

OC Bacteria; Proteobacteria; alpha subdivision; Rhodospirillum;
 OC Rhodospirillum;
 OX NCBI_TaxID=1085;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-R5;
 RX MEDLINE=2038142; PubMed=10671438;
 RA Cheng Y.S., Brantner C.A., Tsapin A., Collins M.L.P.;
 RT "Role of the H protein in assembly of the photochemical reaction
 center and intracytoplasmic membrane in Rhodospirillum rubrum.";
 RL J. Bacteriol. 192:1200-1207 (2000).
 DR EMBL: AF20319; AAF37352.1;
 FT NON-TER
 SQ SEQUENCE 692 AA; 75453 MW; 96430AE93BF35680 CRC64;

Query Match 36.0%; Score 62; DB 2; Length 692;
 Best Local Similarity 40.6%; Pred. No. 0.96;
 Matches 13; Conservative 4; Mismatches 9; Indels 6; Gaps 1;

OY 3 AHQHSMEIRTPDINPAWYAG-----RGIRPV 28
 DB 568 AEOVSLETRTRALNPKWEGMANGVEYRQI 599

RESULT 3
 O97859 PRELIMINARY; PRT; 428 AA.
 AC O97859;
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE GANGLIOSIDE SIALIDASE.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BRAIN;
 RX MEDLINE=99143165; PubMed=9986745;
 RA Miyagi T., Wada T., Iwamatsu A., Hata K., Yoshikawa Y., Tokuyama S.,
 RA Kanada M.;
 RT "Molecular cloning and characterization of a plasma membrane-
 associated sialidase specific for gangliosides.";
 RL J. Biol. Chem. 274:5004-5011 (1999).
 DR EMBL: AB008184; BAA75071.1;
 DR INTERPRO: IPR002860;
 DR Pfam: PF02012; BNR: 3;
 SQ SEQUENCE 428 AA; 47916 MW; 418B34F3245A8F2I CRC64;

Query Match 35.5%; Score 61; DB 6; Length 428;
 Best Local Similarity 40.7%; Pred. No. 0.8;
 Matches 11; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

OY 2 RAHQSMETRPDINPAWYAGRGIRPV 28
 DB 195 RARPHSLMTYSDLGATWHNRLLKPM 221

RESULT 4
 O90UF9 PRELIMINARY; PRT; 54 AA.
 AC O90UF9;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE D44797.3 (SUSHI-REPEAT PROTEIN (SRPUL)) (FRAGMENT).
 GN D44797.3.
 OS Homo sapiens (Human).
 OC Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lawlor S.;
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL035608; CAB55682.1;
 FT NON-TER
 SQ SEQUENCE 54 AA; 6110 MW; E2F3C39F7B961A9F CRC64;

Query Match 33.1%; Score 57; DB 4; Length 54;
 Best Local Similarity 56.2%; Pred. No. 0.34;
 Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 12 TPADINPAWYAGRGIRPV 27
 DB 18 TPATPTWYAGSGYRP 33

RESULT 5
 O60687 PRELIMINARY; PRT; 465 AA.
 AC O60687;
 DT 01-AUG-1998 (TREMBlrel. 07, Created)
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE SRSHT-REPEAT PROTEIN.
 GN SRPUL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kurosawa H., Inukai T., Inaba T., Goi K., Chang K.-S., Sinjo T.,
 RA Kakestrew K.M., Naevø C.W., Look T.A.;
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF060567; AAC15765.1;
 DR INTERPRO: IPR000436;
 DR INTERPRO: IPR001128;
 DR Pfam: PF00084; SUSHI: 3;
 DR PROSITE: PS00086; CYTOCHROME_P450; UNKNOWN_I.
 SQ SEQUENCE 465 AA; 52971 MW; 4D752B187FF3FEB8 CRC64;

Query Match 33.1%; Score 57; DB 4; Length 465;
 Best Local Similarity 56.2%; Pred. No. 0.34;
 Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 12 TPADINPAWYAGRGIRPV 27
 DB 18 TPATPTWYAGSGYRP 33

RESULT 6
 O90U49 PRELIMINARY; PRT; 428 AA.
 AC O90U49;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE GANGLIOSIDE SIALIDASE.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BRAIN;
 RA Wada T., Yoshikawa Y., Tokuyama S., Kuwabara M., Akita H., Miyagi T.,
 RT "Cloning, expression, and chromosomal mapping of a human ganglioside
 sialidase.";

OY 1 SRAHSHMEIRTPDINPAWYAGRGIRPVG 30
 ID 008689 PRELIMINARY: PRT: 176 AA.
 DB 435 SSSASHSVELDTTOAR-EMFMRDVRPEGO 463

RESULT 10
 AC 008689; PRELIMINARY: PRT: 176 AA.
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE CHROMOSOME XV READING FRAME ORF YOR253W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
 OC Saccharomycetaceae; Saccharomycetes.
 NX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MIPS;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RX MEDLINE=97298311; PubMed=9153759;
 RA Jauniaux J.C., Polrey R.;
 RT Sequencing analysis of a 36.8 kb fragment of yeast chromosome XV
 reveals 26 open reading frames including SEC63, COC31, SUG2, GCD1,
 RBL2, PNT1, PAC1 and VPH1.";
 RL Yeast 13:483-487(1997).
 DR EMBL: 275161; CA99475.1; -
 DR INTERPRO: IPR000182; -
 DR INTERPRO: IPR000345; -
 DR PIRAM: PF00583; Acetyltransf. 1.
 DR PROSITE: PS00190; CYTOCHROME C. UNKNOWN 1.
 SQ SEQUENCE 176 AA; 19727 MW; 4F090C597A690BA0 CRC64;

Query Match 31.18; Score 53.5; DB 3; Length 176;
 Best Local Similarity 31.28; Pred. No. 3.8;
 Matches 10; Conservative 5; Mismatches 14; Indels 3; Gaps 1;
 OY 1 SRAHSHMEIRTPDINPAWYAGRGIRPVG 29
 ID 120 SECHQHNVEYVILPAVDLTKRMTANGFEQVG 151

RESULT 11
 AC 09RYP2 PRELIMINARY: PRT: 376 AA.
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DE ADENINE DEAMINASE-RELATED PROTEIN.
 GN DRA0268.
 OS Deinococcus radiodurans.
 OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
 NX NCBI_TaxID=1293;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-RL;
 RX MEDLINE=20036896; PubMed=10567266;
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
 Dodson R.U., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
 Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
 Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,
 Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
 Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
 Fraser C.M.;
 RT "Genome Sequence of the Radioresistant Bacterium Deinococcus
 radiodurans R1.";
 RL Science 286:1571-1577(1999).
 DR EMBL: AEF01863; AAF12376.1; -

DR TIGR: DRA0268; -
 SQ SEQUENCE 376 AA; 39845 MW; 7AB7EF32F8C45651 CRC64;

Query Match 30.88; Score 53; DB 2; Length 376;
 Best Local Similarity 40.08; Pred. No. 9.9;
 Matches 12; Conservative 6; Mismatches 12; Indels 0; Gaps 0;
 OY 1 SRAHSHMEIRTPDINPAWYAGRGIRPVG 30
 ID 177 ARALAHAPDVSRDRBRACAGGRRRPAR 206

RESULT 12
 AC 09ZGE5 PRELIMINARY: PRT: 1292 AA.
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE MG CHELATASE SUBUNIT H BCHH.
 GN BCHH.
 OS Hellobacillus mobilis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Hellobacterium group; Hellobacillus.
 NX NCBI_TaxID=28064;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=99061957; PubMed=9843979;
 RX Xiong J., Inoue K., Bauer C.E.;
 RT "Tracking molecular evolution of photosynthesis by characterization of
 a major photosynthesis gene cluster from Hellobacillus mobilis.";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:14851-14856(1998).
 DR EMBL: AF080002; AAC84033.1; -
 SQ SEQUENCE 1292 AA; 144853 MW; 323AA0517B07448D CRC64;

Query Match 30.28; Score 52; DB 2; Length 1292;
 Best Local Similarity 38.58; Pred. No. 51;
 Matches 10; Conservative 3; Mismatches 7; Indels 6; Gaps 1;
 OY 9 EIRTPDINPAWYAG-----RGIRPV 28
 ID 1176 EPRRTLNPKWTEGMLKHGEGVREI 1201

RESULT 13
 AC 019530 PRELIMINARY: PRT: 294 AA.
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 GN F19H6.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
 OC Rhabditidae; Rhabditinae; Caenorhabditis.
 NX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA McMurray A.;
 RT Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
 RL EMBL: Z50873; CA90762.1; -
 DR EMBL: Z50873; CA90762.1; JOINED.
 DR EMBL: Z68115; CA92169.1; -
 DR EMBL: Z68115; CA92169.1; -
 DR EMBL: Z50873; CA92169.1; JOINED.
 DR HSSP: P24941; ICKP.
 DR INTERPRO: IPR000719; -
 DR INTERPRO: IPR001245; -
 DR INTERPRO: IPR002290; -
 DR PIRAM: PF00069; PKinase; 1.
 DR PRINTS: PRO0109; TYRKINASE.

DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN.1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 SQ SEQUENCE 294 AA: 34006 MW; 071D4664D049FA3 CRC64;

Query Match 29.7%; Score 51; DB 5; Length 294;
 Best Local Similarity 44.8%; Pred. NO. 15;
 Matches 13; Conservative 2; Mismatches 12; Indels 2; Gaps 1;
 QY 3 AHOHSMETPTDINPA--WYAGRGIRPVG 29
 DB 139 AHMHSKRIKMRDIPKPNPFTTNGVLKLG 167

RESULT 14
 ID 087474 PRELIMINARY; PRT; 315 AA.
 AC 087474;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE TRIHYDROXYTOLUENE OXYGENASE.
 GN DNTD.
 OS Burkholderia cepacia (Pseudomonas cepacia).
 OS Bacteria; Proteobacteria; beta subdivision; Burkholderia group;
 OC Burkholderia.
 OX NCBI_TaxID=292;
 RP [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DNT;
 RX MEDLINE=93194809; PubMed=8449889;
 RA Suen W.C., Spain J.C.;
 RT "Cloning and characterization of Pseudomonas sp. strain DNT genes for
 2,4-dinitrotoluene degradation.";
 RL J. Bacteriol. 175:1831-1837(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DNT;
 RA Haigler B.E., Johnson G.R., Suen W.C., Spain J.C.;
 RT "2,4,5-trihydroxytoluene oxygenase catalyzes meta-ring cleavage of
 2,4,5-trihydroxytoluene: Biochemical and genetic evidence.";
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF076848; AAD12738.1;
 DR INTERPRO: IPR000486;
 DR PFAM: PF01013; Extradiol_dioxy; 1.
 DR PRODOM: PD000977; -; 1.
 SQ SEQUENCE 315 AA: 34745 MW; E99261179022961E CRC64;

Query Match 29.7%; Score 51; DB 2; Length 315;
 Best Local Similarity 41.7%; Pred. NO. 16;
 Matches 15; Conservative 3; Mismatches 12; Indels 6; Gaps 2;

QY 1 SRAHSMETPTD--INPAWYAGRGIRP--VGR 30
 DB 216 SRMHGSEFIEDDAQFLGNKMLASRCMKPGWGVGR 251

RESULT 15
 ID 027142 PRELIMINARY; PRT; 430 AA.
 AC 027142;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE CONSERVED PROTEIN.
 GN MTH1070.
 OS Methanobacterium thermoautotrophicum.
 OS Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
 OC Methanobacterium.
 OX NCBI_TaxID=216;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=DELTA H;
 RX MEDLINE=98037514; PubMed=9371463;
 RA Smith D.R., Doucellette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
 RA Aldredge T., Bashirzadeh R., Blakey D., Cook R., Gilbert K.,
 RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
 RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
 RA Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
 RA McDougall S., Shimer G., Goyal A., Pletrowski S., Church G.M.,
 RA Daniels C.J., Mao J.-T., Rice P., Nolling J., Reeve J.N.;
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum
 RT delta: functional analysis and comparative genomics.";
 RL J. Bacteriol. 179:7135-7155(1997).
 DR EMBL: AEO00877; AAB85559.1;
 DR INTERPRO: IPR002510;
 DR PFAM: PF01523; PmbA_TlId; 1.
 SQ SEQUENCE 430 AA: 46062 MW; C4FAE47C111749B6 CRC64;

Query Match 29.7%; Score 51; DB 1; Length 430;
 Best Local Similarity 69.2%; Pred. NO. 22;
 Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 14 DIMPWYAGRGIR 26
 DB 191 DINPEWYAGRGACR 203

Search completed: April 17, 2001, 15:48:03
 Job time: 556 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 17, 2001, 15:39:52 ; Search time 39.1 Seconds

(without alignments)
27.159 Million cell updates/sec

Title: US-09-446-543A-5

Sequence: 1 SRAHSHMELTRPDINPMTAGRGIRPVGRF 31

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	172	100.0	98	1 PRRP_BOVIN	P81264 bos taurus
2	163	94.8	83	1 PRRP_RAT	P81278 rattus norv
3	158	91.9	87	1 PRRP_HUMAN	P18377 homo sapien
4	51	29.7	798	1 UNR_RAT	P18395 rattus norv
5	51	29.7	962	1 UNR_METH	O26543 methanobact
6	50	29.1	973	1 UNR_RHIME	P56899 rhizobium m
7	48.5	28.2	239	1 P6GL_STNY3	P74618 synchocyst
8	48.5	28.2	1882	1 Y468_MYCPN	P75109 mycoplasma
9	48	27.9	424	1 SAM8_SALTY	P23832 salmonella
10	48	27.9	960	1 UNR_TREPA	O00357 treponema p
11	47.5	27.6	1083	1 T2D3_HUMAN	P00268 homo sapien
12	47	27.3	406	1 DH12_RABIT	P51976 oryctolagus
13	47	27.3	972	1 UNR_RABIT	P94572 mycobacteri
14	46.5	27.0	770	1 AVP3_ARATH	P31414 arabidopsis
15	46	26.7	342	1 Y762_METJA	O58172 methanococc
16	46	26.7	347	1 Y576_METJA	O57996 methanococc
17	46	26.7	510	1 YCG8_ECOLI	P29013 escherichia
18	46	26.7	798	1 UNR_HUMAN	O75534 homo sapien
19	46	26.7	836	1 TGK_RABIT	P22758 oryctolagus
20	46	26.7	943	1 UNR_HABIN	P44410 haemophilus
21	45.5	26.5	272	1 TRAZ_DROVI	O02008 drosophila
22	45.5	26.5	311	1 YHHJ_ECOLI	P31993 escherichia
23	45.5	26.5	411	1 PUR7_ARATH	P38025 arabidopsis
24	45.5	26.5	704	1 LIPA_NEITH	O05013 neisseria m
25	45	26.2	482	1 R167_YEAST	P39743 saccharomyc
26	45	26.2	676	1 EXL1_HUMAN	O92935 homo sapien
27	45	26.2	817	1 TGK_RAT	P22735 homo sapien
28	45	26.2	824	1 TGK_RAT	P23606 rattus norv
29	45	26.2	940	1 UNR_ECOLI	P07671 escherichia
30	45	26.2	941	1 UNR_SALTY	P37434 salmonella
31	45	26.2	953	1 UNR_RICPR	O92633 rickettsia
32	45	26.2	1033	1 ACAC_ARATH	O91777 arabidopsis
33	44.5	25.9	152	1 GSTJ_HUMAN	O14880 homo sapien

34	44.5	25.9	241	1 YG3D_YEAST	P53281 saccharomyc
35	44.5	25.9	407	1 Y116_MYCTU	P72052 mycobacteri
36	44.5	25.9	474	1 ATZ6_PSESD	P72156 pseudomonas
37	44.5	25.9	777	1 BISC_ECOLI	P20099 escherichia
38	44.5	25.9	1783	1 Y468_MYCGE	O49460 mycoplasma
39	44	25.6	179	1 RK6_GUTIN	O46908 guillardiella
40	44	25.6	264	1 Y355_BUCAI	P57436 buchiera ap
41	44	25.6	358	1 PLCA_ERMCA	O47499 erwinia chr
42	44	25.6	383	1 CYCR_CHREVI	O82947 chromatiu
43	44	25.6	952	1 UNR_CHREVI	O56242 thermus aqu
44	44	25.6	970	1 UNR_STNY3	P73412 synchocyst
45	44	25.6	1014	1 UNR_STRCO	O92507 streptomyces

ALIGNMENTS

RESULT 1
ID PRRP_BOVIN STANDARD; PRT; 98 AA.
AC P81264;
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PROLACTIN-RELEASING PEPTIDE PRECURSOR (PRP) (PROLACTIN-RELEASING
DE NORMONE) (CONTAINS: PROLACTIN-RELEASING PEPTIDE PRP31; PROLACTIN-
DE RELEASING PEPTIDE PRP20).
GN PRH.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Bovidae; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_Taxid=9913;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-52.
RX MEDLINE=96268781; PubMed=9607765;
RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,
RA Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,
RA Kurokawa T., Nishimura O., Onda H., Fujino M.;
"A prolactin-releasing peptide in the brain."
RA Nature 393:272-276(1998).
CC -!- FUNCTION: STIMULATES PROLACTIN (PR) RELEASE AND REGULATES THE
CC -!- EXPRESSION OF PROLACTIN THROUGH ITS RECEPTOR HGRL. MAY STIMULATE
CC LACTOGENESIS DIRECTLY TO SECRETE PRL.
CC -!- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: AB015417; BAA29025.1;
CC Hormone; Amidation; Signal; Cleavage on pair of basic residues.
CC SIGNAL 1 22
CC PEPTIDE 23 53 PROLACTIN-RELEASING PEPTIDE PRP31.
CC PEPTIDE 33 53 PROLACTIN-RELEASING PEPTIDE PRP20.
CC MOD_RES 53 53 AMIDATION (G-54 PROVIDE AMIDE GROUP).
CC SEQUENCE 98 AA: 10544 MW; 08AC35A13BFA908 CRC64;
CC

Query Match 100.0%; Score 172; DB 1; Length 98;

Best Local Similarity 100.0%; Pred. No. 2.5e-18;

Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 SRAHSHMELTRPDINPMTAGRGIRPVGRF 31
|||||
23 SRAHSHMELTRPDINPMTAGRGIRPVGRF 53

RESULT 2
PRRP_RAT STANDARD: PRT: 83 AA.
AC P81278:
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PROLACTIN-RELEASING PEPTIDE PRECURSOR (PRRP) (PROLACTIN-RELEASING HORMONE) (CONTAINS: PROLACTIN-RELEASING PEPTIDE PRRP31; PROLACTIN-RELEASING PEPTIDE PRRP20).
GN PRH.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RH [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98268781; PubMed=9607765;
RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S., Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M., Kurokawa T., Nishimura O., Onda H., Fujino M.;
RT "A prolactin-releasing peptide in the brain.";
RL Nature 393:272-276(1998).
RL "A prolactin-releasing peptide in the brain.";
CC -1- FUNCTION: STIMULATES PROLACTIN (PRL) RELEASE AND REGULATES THE EXPRESSION OF PROLACTIN THROUGH ITS RECEPTOR HGR3. MAY STIMULATE LACTOGENESIS DIRECTLY TO SECRETE PRL.
CC -1- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AB015418; BAA29026.1; -
KM Hormone; Amidation; Signal; Cleavage on pair of basic residues.
FT SIGNAL 1 21
FT PEPTIDE 22 52 PROLACTIN-RELEASING PEPTIDE PRRP31.
FT PEPTIDE 33 52 PROLACTIN-RELEASING PEPTIDE PRRP20.
FT MOD_RES 52 52 AMIDATION (G-53 PROVIDE AMIDE GROUP).
SQ SEQUENCE 83 AA: 9215 MW: D0C75A264EEB4F29 CRC64;

Query Match 94.8%; Score 163; DB 1; Length 83;
Best Local Similarity 93.5%; Pred. No. 4, 3e-17;
Matches 29; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SRABHSMETRPDIPNPAWYAGRGIRPVGRF 31
DB 22 SRABHSMETRPDIPNPAWYAGRGIRPVGRF 52

RESULT 3
PRRP_HUMAN STANDARD: PRT: 87 AA.
AC P81277:
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PROLACTIN-RELEASING PEPTIDE PRECURSOR (PRRP) (PROLACTIN-RELEASING HORMONE) (CONTAINS: PROLACTIN-RELEASING PEPTIDE PRRP31; PROLACTIN-RELEASING PEPTIDE PRRP20).
GN PRH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RH [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98268781; PubMed=9607765;
RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S., Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M., Kurokawa T., Nishimura O., Onda H., Fujino M.;
RT "A prolactin-releasing peptide in the brain.";
RL Nature 393:272-276(1998).
RL "A prolactin-releasing peptide in the brain.";
CC -1- FUNCTION: STIMULATES PROLACTIN (PRL) RELEASE AND REGULATES THE EXPRESSION OF PROLACTIN THROUGH ITS RECEPTOR HGR3. MAY STIMULATE LACTOGENESIS DIRECTLY TO SECRETE PRL.
CC -1- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X52311; CAA36549.1; -
KM Hormone; Amidation; Signal; Cleavage on pair of basic residues.
FT SIGNAL 1 21
FT PEPTIDE 22 52 PROLACTIN-RELEASING PEPTIDE PRRP31.
FT PEPTIDE 33 52 PROLACTIN-RELEASING PEPTIDE PRRP20.
FT MOD_RES 52 52 AMIDATION (G-53 PROVIDE AMIDE GROUP).
SQ SEQUENCE 87 AA: 9639 MW: 229A2F3F50C9981B CRC64;

Query Match 91.9%; Score 158; DB 1; Length 87;
Best Local Similarity 90.3%; Pred. No. 2, 4e-16;
Matches 28; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SRABHSMETRPDIPNPAWYAGRGIRPVGRF 31
DB 23 SRABHSMETRPDIPNPAWYAGRGIRPVGRF 53

RESULT 4
UHR_RAT STANDARD: PRT: 798 AA.
AC P18395:
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE UHR PROTEIN.
GN UHR.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RH [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=90370473; PubMed=2204029;
RA Jeffers M., Paciucci R., Pellicer A.;
RT "Characterization of uhr, a gene closely linked to H-ras.";
RL Nucleic Acids Res. 18:4891-4893(1990).
CC -1- FUNCTION: RNA-BINDING PROTEIN (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X52311; CAA36549.1; -
KM Hormone; Amidation; Signal; Cleavage on pair of basic residues.
FT SIGNAL 1 21
FT PEPTIDE 22 52 PROLACTIN-RELEASING PEPTIDE PRRP31.
FT PEPTIDE 33 52 PROLACTIN-RELEASING PEPTIDE PRRP20.
FT MOD_RES 52 52 AMIDATION (G-54 PROVIDE AMIDE GROUP).
SQ SEQUENCE 87 AA: 9639 MW: 229A2F3F50C9981B CRC64;

DR	InterPro: IPR002059; -
DR	Pfam: PF00313; CSD; 8.
DR	PROSITE: PS00352; COLD_SHOCK; 4.
DR	KEGG: RNA-binding; Repeat.
FT	DOMAIN 26 87 CSD 1.
FT	DOMAIN 136 179 CSD 2 (INCOMPLETE).
FT	DOMAIN 186 245 CSD 3.
FT	DOMAIN 297 337 CSD 4 (INCOMPLETE).
FT	DOMAIN 349 410 CSD 5.
FT	DOMAIN 447 507 CSD 6.
FT	DOMAIN 519 579 CSD 7.
FT	DOMAIN 610 670 CSD 8.
FT	DOMAIN 674 735 CSD 9.
SO	SEQUENCE 798 AA; 88894 NP; P484B3PABD095A4 CRC64;

Query Match	29.7%	Score 51	DB 1	Length 798
Best Local Similarity	43.5%	Pred. No. 9.2		
Matches 10	Conservative	5	Mismatches	8
			Indels	0
			Gaps	0

Oy 6 HSMERFPDIPNARYAGCIRPV 28

||: | : || | : |||:

Db 583 HSYNCITEFANPTTYSGVIRPL 605

RESULT 5

ID	UVRA_METH	STANDARD	PRT	962 AA
AC	UVRA_METH	STANDARD	PRT	962 AA
AD	026543			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	EXCINUCLEASE ABC SUBUNIT A.			
GN	UVRA OR MTH443.			
OS	Methanobacterium thermoautotrophicum.			
OC	Archaea: Euryarchaeota; Methanobacteriales; Methanobacteriaceae;			
OC	Methanobacterium.			
OX	NCBI_TaxID=145262;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-DELTA H;			
RX	MEDLINE=98037514; PubMed=9371463;			
RA	Smith D R., Doncotte-Stamm L.A., Delouphery C., Lee H.-M., Dubois J.,			
RA	Alredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,			
RA	Harrison D., Hoang L., Keagle P., Lumu W., Pothier B., Qiu D.,			
RA	Spedafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,			
RA	Jivani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,			
RA	McDougal S., Shimer G., Goyal A., Pietrowski S., Church G.M.,			
RA	Daniels C.J., Mac J.-I., Rice P., Nolling J., Reeve J.N.;			
RT	*Complete genome sequence of Methanobacterium thermoautotrophicum			
RL	delta: functional analysis and comparative genomics.";			
RL	J. Bacteriol. 179:7135-7152(1997).			
CC	-1 FUNCTION: THE ABC EXCISION NUCLEASE IS A DNA REPAIR ENZYME THAT			
CC	CATALYZES THE EXCISION REACTION OF UV-DAMAGED NUCLEOTIDE SEGMENTS			
CC	PRODUCING OLIGOMERS HAVING THE MODIFIED BASE(S). UVRA IS AN APASE			
CC	AND A DNA-BINDING PROTEIN THAT PREFERENTIALLY BINDS SINGLE-			
CC	STANDED OR UV-IRRADIATED DOUBLE-STANDED DNA (BY SIMILARITY).			
CC	-1 SUBUNIT: CONSISTS OF THREE SUBUNTS: UVRA, UVRA AND UVRC.			
CC	-1 SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).			
CC	-1 SIMILARITY: BELONGS TO THE ABC-BINDING TRANSPORT PROTEIN FAMILY			
CC	(ABC TRANSPORTERS). CONTAINS TWO ABC DOMAINS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation - its			
CC	the European Bioinformatics Institute. There are no restrictions on			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL: AE000828; AAB84949.1; ..			
DR	InterPro: IPR001617; ..			
DR	pfam: PF000005; ABC_tran. 2.			

DR	PROSITE; AS00211; ABC_TRANSPORTER; 2.
KM	DNA response; Excision nuclease; DNA repair; ATP-binding; Regatt;
FM	DNA-binding; zinc-finger.
FT	NP_BIND 38 ATP (POTENTIAL).
FT	NP_BIND 649 ATP (POTENTIAL).
FT	NP_BIND 748 C4 TYPE.
FT	SEQUENCE 962 AA; 108395 MW; 2C0EE7FEC41CDD060 CRC64;

```

Query Match          29.7%; Score 51; DB 1; Length 962;
Best Local Similarity 39.5%; Pred. No. 11;
Matches 15; Conservative 0; Mismatches 5; Indels 18; Gaps 2;

OY      11 RTPDINPAMYAG-----KGRVGRF 31.
      11 11 11 11 11 11 11 11 11 11
Db      703 RTPRNPATYGVFTTHRELPQTPEARKNGRP-GRF 739

RESULT 6
ID      UVRA_RHIME      STANDARD;      PRT;      973 AA.
AC      P56899;
DT      30-MAY-2000 (Ref. 39, Last sequence update)
DT      30-MAY-2000 (Ref. 39, Last sequence update)
DE      EXCINUCLEASE ABC SUBUNIT A.
GN      UVRA.
OS      Rhizobium meliloti (Sinorhizobium meliloti).
OC      Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC      Rhizobiaceae; Sinorhizobium.
OX      NCBI_TaxID=382;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN-1021;
RA      Gilbert F., Capela D., Hubler-Barloy F., Gatus M., Batut J.,
RA      Bolstad F., Gouzy J., Kahn D., Thebaud P., Goffeau A.,
RA      Purnelle B., Pohl T., Bolne G., Schneider S., Portetelle D.,
RA      Vandemool M., Puehler A., Becker A., Weidner S.;
RA      Submitted (May-2000) to the SWISS-PROT data bank.
RN      [2]
RP      SEQUENCE OF 1-140 FROM N.A.
RC      STRAIN-2021;
RX      MEDLINE=93430868; PubMed=10503543;
RT      Tapias A., Barbe J.;
RT      "Regulation of divergent transcription from the uvra-ssb promoters in
RT      Sinorhizobium meliloti.";
RL      Mol. Gen. Genet. 262:121-130(1999).
CC      -1- FUNCTION: THE ABC EXCISION NUCLEASE IS A DNA REPAIR ENZYME THAT
CC      CATALYZES THE EXCISION REACTION OF UV-DAMAGED NUCLEOTIDE SEGMENTS
CC      PRODUCING OLIGOMERS HAVING THE MODIFIED BASE(S). UVRA IS AN ATPASE
CC      AND A DNA-BINDING PROTEIN THAT PREFERENTIALLY BINDS SINGLE-
CC      STANDED OR UV-IRRADIATED DOUBLE-STANDED DNA (BY SIMILARITY).
CC      -1- SUBUNIT: CONSISTS OF THREE SUBUNITS: UVRA, UVRB AND UVRC.
CC      -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC      -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC      (ABC TRANSPORTERS). CONTAINS TWO ABC DOMAINS.
CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC
CC      EMBL: AF125162; AAF03210.1; -.
CC      InterPro: IPR001617; -.
KM      PROSITE: PS00211; ABC_TRANSPORTER; PARTIAL.
KM      SOS response: Excision nuclease; DNA repair; ATP-binding; Repeat;
KM      DNA-binding; Zinc-finger.
FT      NP_BIND      34      41      ATP (POTENTIAL).
FT      NP_BIND      662    669      ATP (POTENTIAL).
FT      ZN_FING      761      787      C4-TYPE.

```



```

SAMB_SALTY
ID SAMB_SALTY STANDARD; PRT; 424 AA.
AC P23832;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE SAMB PROTEIN.
GN SAMB.
OS Salmonella typhimurium.
OC Plasmid 60-Mda cryptic.
CC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
CC Salmonella.
CC NCBI_TaxID=602;
CC 11
CC SEQUENCE FROM N.A.
CC STRAIN-LET2;
CC MEDLINE=91123176; PubMed=1991707;
CC Mohani T., Hakura A., Nakai Y., Watanabe M., Murayama S.Y.,
CC Sotani T.;
CC "Salmonella typhimurium has two homologous but different umuDC
CC operators: cloning of a new umuOC-like operon (samb) present in a
CC 60-megadalton cryptic plasmid of S. typhimurium."
CC J. Bacteriol. 173:1051-1063(1991).
CC or send an email to license@lsb-slb.ch.
CC -----
CC -1- SIMILARITY: BELONGS TO THE IMPB/MDCB/SAMB FAMILY.
CC
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.lsb-slb.ch/announce/
CC or send an email to license@lsb-slb.ch).
CC -----
CC DR EMBL: D90502; BAAL1426.1; -
CC DR PIR: B38176; B38176.
CC DR InterPro: IPR001126; -
CC DR Pfam: PF00817; IMS; 1.
CC KM Plasmid: SOS mutagenesis; DNA repair.
CC KW SEQUENCE 424 AA; 47727 MW; FF8CA476CC58A2B CRC64;
CC
Query Match 27.9%; Score 48; DB 1; Length 424;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 9; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

```

```

RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete."
RL Science 281:375-388(1998).
CC -1- FUNCTION: THE ABC EXCISION NUCLEASE IS A DNA REPAIR ENZYME THAT
CC CATALYZES THE EXCISION REACTION OF UV-DAMAGED NUCLEOTIDE SEGMENTS
CC PRODUCING OLIGOMERS HAVING THE MODIFIED BASE(S). UVRA IS AN ATPASE
CC AND A DNA-BINDING PROTEIN THAT PREFERENTIALLY BINDS SINGLE-
CC STRANDED OR UV-IRRADIATED DOUBLE-STRANDED DNA (BY SIMILARITY).
CC -1- SUBUNIT: CONSISTS OF THREE SUBUNITS; UVRA, UVRA AND UVRC.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC (ABC TRANSPORTERS). CONTAINS TWO ABC DOMAINS.
CC
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.lsb-slb.ch/announce/
CC or send an email to license@lsb-slb.ch).
CC -----
CC DR EMBL: AE001227; AAC65502.1; -
CC DR TIGR: TP0514; -
CC DR InterPro: IPR001617; -
CC DR Pfam: PF00005; ABC_tran; 2.
CC DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
CC KW SOS response; Excision nuclease; DNA repair; ATP-binding; Repeat;
CC DNA-binding; Zinc-finger.
CC NP-BIND 35 42
CC NP-BIND 657 664 ATP (POTENTIAL).
CC FT NP-BIND 270 297 ATP (POTENTIAL).
CC FT ZN-FING 756 782 CA-TYPE.
CC FT ZN-FING 756 782 CA-TYPE.
CC SEQUENCE 960 AA; 106010 MW; 32F78624B19F7ABF CRC64;
CC
Query Match 27.9%; Score 48; DB 1; Length 960;
Best Local Similarity 36.8%; Pred. No. 31;
Matches 14; Conservative 1; Mismatches 5; Indels 18; Gaps 2;

```

```

QY 10 ITPDINPARYAGKIRP 27
ID 1 1 1 1 1 1 1 1
DB 381 INHPEKGRWFGAGRIAP 398

```

```

QY 11 RFPDINPARYAG-----RGIRPVGRF 31
ID 1 1 1 1 1 1 1 1
DB 711 RFPDINPARYAGVETDIRMLFSGVPEAKMKGRYP-GRF 747

```

```

RESULT 10
ID UVRA_TREPA STANDARD; PRT; 960 AA.
AC 083527;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE EXCINUCLEASE ABC SUBUNIT A.
GN UVRA OR TP0514.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
CC NCBI_TaxID=160;
CC 11
CC SEQUENCE FROM N.A.
CC STRAIN=NICHOLS;
CC MEDLINE=98332770; PubMed=9665876;
CC Fraser C.M., Norris S.J., Weinstein G.M., White O., Sutton G.G.,
CC Dodson R., Gwin M., Hickey E.K., Clayton R., Kechum K.A.,
CC Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
CC Khalak H., Richardson D., Howell J.K., Chidambaram M., Uterback T.,
CC McDonald L., Artach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
CC Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,

```

```

RESULT 11
ID T2D3_HUMAN STANDARD; PRT; 1083 AA.
AC 000268; Q99721;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE TRANSCRIPTION INITIATION FACTOR TFIID 135 KDa SUBUNIT (TAFI1-135)
DE (TAFI1135) (TAFI1-130) (TAFI1130).
GN TAF2C1 OR TAF2C OR TAFI1135 OR TAFI1130.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
CC NCBI_TaxID=9606;
CC 11
CC SEQUENCE FROM N.A.
CC MEDLINE=97336072; PubMed=9192867;
CC Mengus G., May M., Chamber P., Davidson I.;
CC "Human TAF(II)135 potentiates transcriptional activation by the AF-2s
CC of the retinoic acid, vitamin D3, and thyroid hormone receptors in
CC RT mammalian cells."
CC RT Genes Dev. 11:1381-1395(1997).
CC 12
CC SEQUENCE OF 105-1083 FROM N.A. AND PARTIAL SEQUENCE.
CC MEDLINE=97098442; PubMed=8942982;
CC Tanase N., Saluja D., Vassallo M.F., Chen J.-L., Admon A.;
CC "Molecular cloning and analysis of two subunits of the human TFIID
CC complex: hTAF1130 and hTAF1100."

```

CC Proc. Natl. Acad. Sci. U.S.A. 93:13611-13616(1996).

CC -1- FUNCTION: MAKES PART OF TRIID IS A MUTIMERIC PROTEIN COMPLEX THAT

CC PLAYS A CENTRAL ROLE IN MEDIATING PROMOTER RESPONSES TO VARIOUS

CC ACTIVATORS AND REPRESSORS. POTENTIATES TRANSCRIPTIONAL ACTIVATION

CC BY THE AF-2S OF THE RETINOIC ACID, VITAMIN D3 AND THYROID HORMONE.

CC -1- SUBUNIT: TRIID IS COMPOSED OF TATA BINDING PROTEIN (TBP) AND A

CC NUMBER OF TBP-ASSOCIATED FACTORS (TAFs).

CC -1- SUBCELLULAR LOCATION: NUCLEAR

CC -1- SIMILARITY: BELONGS TO THE TAF2C FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC DR EMBL: Y11354; CAA2189.1; -

CC DR EMBL: U75308; AAC50901.1; -

CC DR HSSP: P19656; IAFH.

CC DR MIN: 601796; -

CC KW Transcription regulation; Nuclear protein.

CC FT DOMAIN 39 42 POLY-HIS.

CC FT DOMAIN 52 57 POLY-ALA.

CC FT DOMAIN 98 101 POLY-GLY.

CC FT DOMAIN 142 148 POLY-ALA.

CC FT DOMAIN 268 275 POLY-PRO.

CC FT DOMAIN 331 337 POLY-ALA.

CC FT DOMAIN 680 683 POLY-PRO.

CC FT DOMAIN 808 813 POLY-ASP.

CC FT DOMAIN 828 831 POLY-ASP.

CC FT CONFLICT 105 117 PGPPRRPRIPVA -> GRGLLQGRGREGS

CC FT CONFLICT 233 264 MISSING (IN REF. 2).

CC FT CONFLICT 293 293 P -> L (IN REF. 2).

CC SQ SEQUENCE 1083 AA; 109943 MW; A6453827572A0752 CRC64;

Query Match 27.6%; Score 47.5; DB 1; Length 1083;

Best Local Similarity 30.8%; Pred. No. 41;

Matches 12; Conservative 4; Mismatches 8; Indels 15; Gaps 1;

OY 1 SRANQSMKIRTPDI-----NPAWYAGRG 24

DB 36 SAANHNLAPRTPYRAAAGLGNHYVSGSAGAGAG 74

RESULT 12

ID DH12_RABIT STANDARD; PRT; 406 AA.

AC P51976:

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 01-OCT-1996 (Rel. 34, Last annotation update)

DE CORICOSTEROID 11-BETA-DEHYDROGENASE, ISOZYME 2 (EC 1.1.1.146) (11-

DE DH2) (11-BETA-NYDROXYSTEROID DEHYDROGENASE 2) (11-BETA-HSD2) (NAD-

DE DEPENDENT 11-BETA-NYDROXYSTEROID DEHYDROGENASE).

GN NSD1B2.

OS Oryctolagus cuniculus (Rabbit).

CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

CC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

OX NCBI_TaxID=9986;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=NEW ZEALAND WHITE; TISSUE=Kidney cortex;

RA MEDLINE=95269702; PubMed=7750480;

RA Naray-Pejes-Toth A., Fejes-Toth G.;

RT "Expression cloning of the aldosterone target cell-specific 11 beta-

RT hydroxysteroid dehydrogenase from rabbit collecting duct cells.";

RL Endocrinology 136:2579-2586(1995).

CC -1- FUNCTION: HAS A ROLE IN MODULATING GLUCOCORTICOID ACTIVITY BOTH AT

CC THE LEVEL OF THE MINERALCORTICOID RECEPTOR AND THE GLUCOCORTICOID

CC RECEPTOR. USES NADH WHILE 11-DH1 USES NADPH. CATALYZES NON

CC REVERSIBLE THE CONVERSION OF CORTISOL TO THE INACTIVE METABOLITE

CC CORTISONE (BY SIMILARITY).

CC -1- CATALYTIC ACTIVITY: A 11-BETA-HYDROXYSTEROID + NAD(+) -> A

CC 11-OXOSTEROID + NADH.

CC -1- SUBCELLULAR LOCATION: MICROSOAL.

CC -1- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN THE KIDNEY.

CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES

CC (SDR) FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC DR EMBL: U23835; AAA86387.1; -

CC DR HSSP: P14061; IFDV.

CC DR InterPro: IPR002198; -

CC DR Pfam: PF00106; adh_short.1.

CC DR PROSITE: PS00061; ADN_SHORT.1.

CC KM Oxidoreductase; NAD; Microsome.

CC FT NR_BIND 82 111 NAD (BY SIMILARITY).

CC FT ACT_SITE 232 232 BY SIMILARITY.

CC SQ SEQUENCE 406 AA; 44181 MW; F14579663918425D CRC64;

Query Match 27.3%; Score 47; DB 1; Length 406;

Best Local Similarity 25.0%; Pred. No. 18;

Matches 10; Conservative 6; Mismatches 8; Indels 16; Gaps 1;

OY 2 RAHQSMKIRTPDINPA-----WYAGRG 25

DB 305 REFLHSRLALPDISPVADITALLAAPPRTYRGRL 344

RESULT 13

ID UYRA_MYCTU STANDARD; PRT; 972 AA.

AC P94972:

DT 15-DEC-1998 (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE EXONUCLEASE ABC SUBUNIT A.

GN UYRA OR RV1638 OR MTCY06M11.02.

OS Mycobacterium tuberculosis.

CC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

CC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.

OX NCBI_TaxID=1773;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=H37RV;

RA MEDLINE=98295987; PubMed=9634230;

RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,

RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekela F.,

RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,

RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,

RA Hornsby F., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,

RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,

RA Rutter S., Seeger K., Skellern S., Squares S., Squires R., Sulston J.E.,

RA Taylor K., Whitehead S., Barrett B.G.;

RT "Deciphering the biology of Mycobacterium tuberculosis from the

RT complete genome sequence.";

RL Nature 393:537-544(1998).

CC -1- FUNCTION: THE ABC EXCISION NUCLEASE IS A DNA REPAIR ENZYME THAT

CC CATALYZES THE EXCISION REACTION OF UV-DAMAGED NUCLEOTIDE SEGMENTS

CC PRODUCING OLIGOMERS HAVING THE MODIFIED BASE(S). UYRA IS AN ATPASE

CC AND A DNA-BINDING PROTEIN THAT PREFERENTIALLY BINDS SINGLE-

CC STRANDED OR UV-IRRADIATED DOUBLE-STRANDED DNA (BY SIMILARITY).

CC -1- SUBUNIT: CONSISTS OF THREE SUBUNITS; UYRA, UYRB AND UYRC.

CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
 CC (ABC TRANSPORTERS). CONTAINS TWO ABC DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Z85982; CAB0633.1; -.
 DR Uniref: K1638; -.
 DR InterPro: IPR001617; -.
 DR Pfam: PF00005; ABC_tran; 2.
 DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
 KW SOS response: Exclusion nuclease; DNA repair; ATP-binding; Repeat;
 FW DNA-binding; Zinc-finger.
 FT NP-BIND 32 39 ATP (POTENTIAL).
 FT NP-BIND 654 661 ATP (POTENTIAL).
 FT ZN_FING 257 285 C4-TYPE (ATYPICAL).
 FT ZN_FING 753 779 C4-TYPE.
 SQ SEQUENCE 972 AA; 106131 MW; 8937A764E592D981 CRC64;
 Query Match 27.3%; Score: 47; DB 1; Length 972;
 Best Local Similarity 36.8%; Pred. Ho. 43;
 Matches 14; Conservative 1; Mismatches 5; Indels 18; Gaps 2;
 Oy 11 RTDINPAMVAG-----NGIRPVGRF 31
 Db 708 RTPSHRATYGVFKIRLTLEAATKAVGYGP-GRF 744
 RESULT 14
 AVP3_ARATH STANDARD; PRT; 770 AA.
 AC P31414; Q41919; -.
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE PYROPHOSPHATE-ENERGIZED VACUOLAR MEMBRANE PROTON PUMP (EC 3.6.1.1)
 GN (PYROPHOSPHATE-ENERGIZED INORGANIC PYROPHOSPHATASE) (H+-PPASE).
 OS AWP-3.
 OC Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RH [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RC MEDLINE=92179265; PubMed=1311852;
 RA Sarafian V., Kim Y., Poole R.J., Rea P.A.;
 RT "Molecular cloning and sequence of cDNA encoding the pyrophosphate-
 RT energized vacuolar membrane proton pump of Arabidopsis thaliana.";
 RT Proc. Natl. Acad. Sci. U.S.A. 89:1775-1779(1992).
 RL [2]
 RP SEQUENCE OF 554-671 FROM H.A.
 RC STRAIN-CV. COLUMBIA; TISSUE=Green siliques;
 RA Raynal M., Grellier F., Laude M., Meyer Y., Cooke R., Delseny M.;
 RL Submitted (OCT-1992) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: CONTRIBUTES TO THE TRANSTONOLAST (FROM CYTOSOL TO
 CC VACUOLE LUMEN) H+-ELECTROCHEMICAL POTENTIAL DIFFERENCE. THE
 CC MAGNITUDE THAT THE H+-PPASE ON THE SAME MEMBRANE.
 CC -1- CATALYTIC ACTIVITY: PYROPHOSPHATE + H(2)O -> 2 ORTHOPHOSPHATE.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. VACUOLAR
 CC (TONOPLAST).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M81892; AAA32754.1; -.
 DR EMBL: Z17694; CAA19038.1; -.
 DR PIR: A38230; A38230.
 KW Hydrogen ion transport; Hydrolase; Transmembrane; Magnesium.
 FT DOMAIN 1 13 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 14 34 INTRAVACUOLAR (POTENTIAL).
 FT TRANSMEM 35 100 POTENTIAL.
 FT TRANSMEM 101 121 POTENTIAL.
 FT TRANSMEM 122 142 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 143 163 POTENTIAL.
 FT TRANSMEM 164 195 INTRAVACUOLAR (POTENTIAL).
 FT TRANSMEM 196 216 POTENTIAL.
 FT TRANSMEM 217 327 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 328 348 POTENTIAL.
 FT TRANSMEM 349 368 INTRAVACUOLAR (POTENTIAL).
 FT TRANSMEM 369 389 POTENTIAL.
 FT TRANSMEM 390 405 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 406 426 POTENTIAL.
 FT TRANSMEM 427 484 INTRAVACUOLAR (POTENTIAL).
 FT TRANSMEM 485 505 POTENTIAL.
 FT TRANSMEM 506 542 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 543 563 POTENTIAL.
 FT TRANSMEM 564 573 INTRAVACUOLAR (POTENTIAL).
 FT TRANSMEM 574 594 POTENTIAL.
 FT TRANSMEM 595 645 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 646 666 POTENTIAL.
 FT TRANSMEM 667 687 INTRAVACUOLAR (POTENTIAL).
 FT TRANSMEM 688 688 POTENTIAL.
 FT TRANSMEM 689 746 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 747 767 POTENTIAL.
 FT TRANSMEM 768 770 INTRAVACUOLAR (POTENTIAL).
 FT TRANSMEM 771 770 T -> A (IN REF. 2).
 FT TRANSMEM 771 770 L -> P (IN REF. 2).
 FT TRANSMEM 771 770 L -> P (IN REF. 2).
 SQ SEQUENCE 770 AA; 80819 MW; CE713B42B299860 CRC64;
 Query Match 27.0%; Score: 46.5; DB 1; Length 770;
 Best Local Similarity 26.3%; Pred. No. 40;
 Matches 10; Conservative 9; Mismatches 10; Indels 9; Gaps 1;
 Oy 1 SRAHSHSEIRTPDI-----NPATYAGRGIRPVG 29
 Db 565 SRAGIHVDVLTIPKYLIGLVGAMLPWFSAWTKRVS 602
 RESULT 15
 Y762_METUA STANDARD; PRT; 342 AA.
 ID Y762_METUA
 AC Q58172;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE HYPOTHETICAL PROTEIN M0762.
 GN M0762.
 OS Methanococcus jannaschii.
 OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
 OC Methanococcus.
 OX NCBI_TaxID=2190;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
 RC MEDLINE=96337999; PubMed=8688087;
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kleravage A.R., Dougherty B.A., Tomb J.F., Adams M.D., Reich C.I.,
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glöck A.,
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 17, 2001, 15:34:31 : Search time 61.54 Seconds
(without alignments)
9.677 Million cell updates/sec

Title: US-09-446-543A-5

Perfect score: 172

Sequence: 1 SRAHQSMETRPDIPNPMYAGKIRPVRE 31

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/prodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCRTS.COMB.pep:*
6: /cgn2_6/prodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	172	100.0	31	3	US-09-105-678A-7
2	172	100.0	31	3	US-09-105-678A-31
3	172	100.0	32	3	US-09-105-678A-32
4	172	100.0	33	3	US-09-105-678A-33
5	163	94.8	31	3	US-09-105-678A-8
6	163	94.8	31	3	US-09-105-678A-37
7	163	94.8	31	4	US-09-172-353-4
8	163	94.8	32	3	US-09-105-678A-38
9	163	94.8	33	3	US-09-105-678A-39
10	161	93.6	29	3	US-09-105-678A-29
11	158	91.9	31	3	US-09-105-678A-43
12	158	91.9	32	3	US-09-105-678A-44
13	158	91.9	33	3	US-09-105-678A-45
14	158	91.9	30	3	US-09-105-678A-46
15	158	91.9	23	3	US-09-105-678A-34
16	158	91.9	22	3	US-09-105-678A-35
17	158	91.9	21	3	US-09-105-678A-36
18	158	91.9	20	3	US-09-105-678A-40
19	158	91.9	21	3	US-09-105-678A-41
20	158	91.9	22	3	US-09-105-678A-42
21	158	91.9	23	3	US-09-105-678A-43
22	158	91.9	24	3	US-09-105-678A-44
23	158	91.9	25	3	US-09-105-678A-45
24	158	91.9	26	3	US-09-105-678A-46
25	158	91.9	27	3	US-09-105-678A-47
26	158	91.9	28	3	US-09-105-678A-48
27	158	91.9	29	3	US-09-105-678A-49

28	51	29.7	239	3	US-09-295-068-3	Sequence 3, Appl
29	51	29.7	239	3	US-09-111-444-7	Sequence 7, Appl
30	50	29.1	349	5	US-08-118-270-71	Sequence 71, Appl
31	50	29.1	349	5	PCT-US93-08528-71	Sequence 71, Appl
32	46	26.7	555	2	US-08-982-232-7	Sequence 7, Appl
33	46	26.7	555	2	US-08-982-232-14	Sequence 14, Appl
34	46	26.7	774	3	US-08-902-632-2	Sequence 2, Appl
35	46	26.7	774	3	US-09-073-354-1	Sequence 1, Appl
36	46	26.7	774	3	US-08-656-005A-1	Sequence 1, Appl
37	46	26.7	774	4	US-09-073-259-1	Sequence 1, Appl
38	46	26.7	774	4	US-09-363-095-1	Sequence 1, Appl
39	45	26.2	440	3	US-08-985-335-3	Sequence 3, Appl
40	45	26.2	940	4	US-09-078-347A-1	Sequence 1, Appl
41	44.5	25.9	152	2	US-08-959-865-1	Sequence 1, Appl
42	44	25.6	162	4	US-08-992-176-9	Sequence 9, Appl
43	44	25.6	4551	4	US-09-320-878-1	Sequence 1, Appl
44	43	25.0	386	2	US-08-663-310-6	Sequence 6, Appl
45	43	25.0	386	2	US-09-006-491-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-09-105-678A-7
Sequence 7, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESS: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-7

Query Match 100.0%, Score 172, DB 3, Length 31,
Best local Similarity 100.0%, Pred. No. 6.6e-20,
Matches 31, Conservative 0, Mismatches 0, Indels 0, Gaps 0,

OY 1 SRAHSHMEIRTPDINPAMYAGRGIRPVGRF 31
DB 1 SRAHSHMEIRTPDINPAMYAGRGIRPVGRF 31

RESULT 2

US-09-105-678A-31
Sequence 31, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEO ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-31

Query Match 100.0%; Score 172; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 6,6e-20;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRAHSHMEIRTPDINPAMYAGRGIRPVGRF 31
DB 1 SRAHSHMEIRTPDINPAMYAGRGIRPVGRF 31

RESULT 3

US-09-105-678A-32
Sequence 32, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEO ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 32 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-32

Query Match 100.0%; Score 172; DB 3; Length 32;
Best Local Similarity 100.0%; Pred. No. 6,9e-20;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRAHSHMEIRTPDINPAMYAGRGIRPVGRF 31
DB 1 SRAHSHMEIRTPDINPAMYAGRGIRPVGRF 31

RESULT 4

US-09-105-678A-33
Sequence 33, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.

REGISTRATION NUMBER: 27, 026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-33

Query Match 100.0%; Score 172; DB 3; Length 33;
Best Local Similarity 100.0%; Pred. No. 7.2e-20;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SRAHSHMETRTPDINPAMYAGRGIRPVGRF 31
DB 1 SRAHSHMETRTPDINPAMYAGRGIRPVGRF 31

RESULT 5
US-09-105-678A-8
Sequence 8, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105, 678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27, 026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-8

Query Match 94.8%; Score 163; DB 3; Length 31;
Best Local Similarity 93.5%; Pred. No. 1.6e-18;
Matches 29; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SRAHSHMETRTPDINPAMYAGRGIRPVGRF 31
DB 1 SRAHSHMETRTPDINPAMYAGRGIRPVGRF 31

RESULT 6
US-09-105-678A-37
Sequence 37, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105, 678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27, 026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-37

Query Match 94.8%; Score 163; DB 3; Length 31;
Best Local Similarity 93.5%; Pred. No. 1.6e-18;
Matches 29; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SRAHSHMETRTPDINPAMYAGRGIRPVGRF 31
DB 1 SRAHSHMETRTPDINPAMYAGRGIRPVGRF 31

RESULT 7
US-09-172-353-4
Sequence 4, Application US/09172353
Patent No. 6197530
GENERAL INFORMATION:
APPLICANT: Stricker-Kongre, Alain
APPLICANT: Gu, Wei
TITLE OF INVENTION: GP10 AS A TARGET FOR IDENTIFYING WEIGHT MODULATING COMPOUNDS
FILE REFERENCE: 07334/102001
CURRENT APPLICATION NUMBER: US/09/172, 353
CURRENT FILING DATE: 1998-10-14
NUMBER OF SEQ ID NOS: 7

SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 31
; TYPE: prt
; ORGANISM: Mus musculus
US-09-172-353-4

Query Match 94.8%; Score 163; DB 4; Length 31;
Best local Similarity 93.5%; Pred. No. 1.6e-18;
Matches 29; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 SRAHSHMETRPDINPAMYAGRGIRPVGRF 31
|||||
DB 1 SRAHSHMETRPDINPAMYTGGRIRPVGRF 31

RESULT 8
US-09-105-678A-38

; Sequence 38, Application US/09105678A
; Patent No. 6103882

; GENERAL INFORMATION:

; APPLICANT: Suenaga, Masato

; APPLICANT: Moriya, Takeo

; APPLICANT: Nishimura, Osamu

; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND

; NUMBER OF SEQUENCES: 52

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

; STREET: 130 Water Street

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/105,678A

; FILING DATE: 26-JUN-1998

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: JP 172118/1997

; FILING DATE: 27-JUN-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Conlin, David G.

; REGISTRATION NUMBER: 27,026

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-523-3400

; TELEFAX: 617-523-6440

; INFORMATION FOR SEQ ID NO: 38:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 32 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; US-09-105-678A-38

Query Match 94.8%; Score 163; DB 3; Length 32;
Best local Similarity 93.5%; Pred. No. 1.6e-18;
Matches 29; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 SRAHSHMETRPDINPAMYAGRGIRPVGRF 31
|||||
DB 1 SRAHSHMETRPDINPAMYTGGRIRPVGRF 31

RESULT 9

US-09-105-678A-39
; Sequence 39, Application US/09105678A
; Patent No. 6103882

; GENERAL INFORMATION:

; APPLICANT: Suenaga, Masato

; APPLICANT: Moriya, Takeo

; APPLICANT: Nishimura, Osamu

; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND

; NUMBER OF SEQUENCES: 52

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

; STREET: 130 Water Street

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/105,678A

; FILING DATE: 26-JUN-1998

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: JP 172118/1997

; FILING DATE: 27-JUN-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Conlin, David G.

; REGISTRATION NUMBER: 27,026

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-523-3400

; TELEFAX: 617-523-6440

; INFORMATION FOR SEQ ID NO: 39:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 33 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; US-09-105-678A-39

Query Match 94.8%; Score 163; DB 3; Length 33;
Best local Similarity 93.5%; Pred. No. 1.7e-18;
Matches 29; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 SRAHSHMETRPDINPAMYAGRGIRPVGRF 31
|||||
DB 1 SRAHSHMETRPDINPAMYTGGRIRPVGRF 31

RESULT 10

; Sequence 29, Application US/09105678A

; Patent No. 6103882

; GENERAL INFORMATION:

; APPLICANT: Suenaga, Masato

; APPLICANT: Moriya, Takeo

; APPLICANT: Nishimura, Osamu

; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND

; NUMBER OF SEQUENCES: 52

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

; STREET: 130 Water Street

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 29 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-29

Query Match 93.6%; Score 161; DB 3; Length 29;
Best Local Similarity 100.0%; Pred. No. 3e-18;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SRAHSHMEIRTPDINPAMYAGRIPIVGR 29
Db 1 SRAHSHMEIRTPDINPAMYAGRIPIVGR 29

RESULT 11
US-09-105-678A-9
Sequence 9, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSER: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-9

Query Match 91.9%; Score 158; DB 3; Length 31;
Best Local Similarity 90.3%; Pred. No. 9.2e-18;
Matches 28; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Qy 1 SRAHSHMEIRTPDINPAMYAGRIPIVGR 31
Db 1 SRAHSHMEIRTPDINPAMYAGRIPIVGR 31

RESULT 12
US-09-105-678A-43
Sequence 43, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSER: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-43

Query Match 91.9%; Score 158; DB 3; Length 31;
Best Local Similarity 90.3%; Pred. No. 9.2e-18;
Matches 28; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Qy 1 SRAHSHMEIRTPDINPAMYAGRIPIVGR 31
Db 1 SRAHSHMEIRTPDINPAMYAGRIPIVGR 31

```

RESULT 13
US-09-105-678A-44
; Sequence 44, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,678A
; FILING DATE: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-105-678A-44

Query Match          91.9%; Score 158; DB 3; Length 32;
Best Local Similarity 90.3%; Pred. No. 9,6e-18;
Matches 28; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SRAHQSMETRTPDINPANYAGRGIRPVGRF 31
DB 1 STHRHSMEIRTPDINPANYASRGIRPVGRF 31

RESULT 14
US-09-105-678A-45
; Sequence 45, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,678A
; FILING DATE: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-105-678A-45

```

```

Query Match          91.9%; Score 158; DB 3; Length 33;
Best Local Similarity 90.3%; Pred. No. 1e-17;
Matches 28; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 SRAHQSMETRTPDINPANYAGRGIRPVGRF 31
DB 1 STHRHSMEIRTPDINPANYASRGIRPVGRF 31

```

```

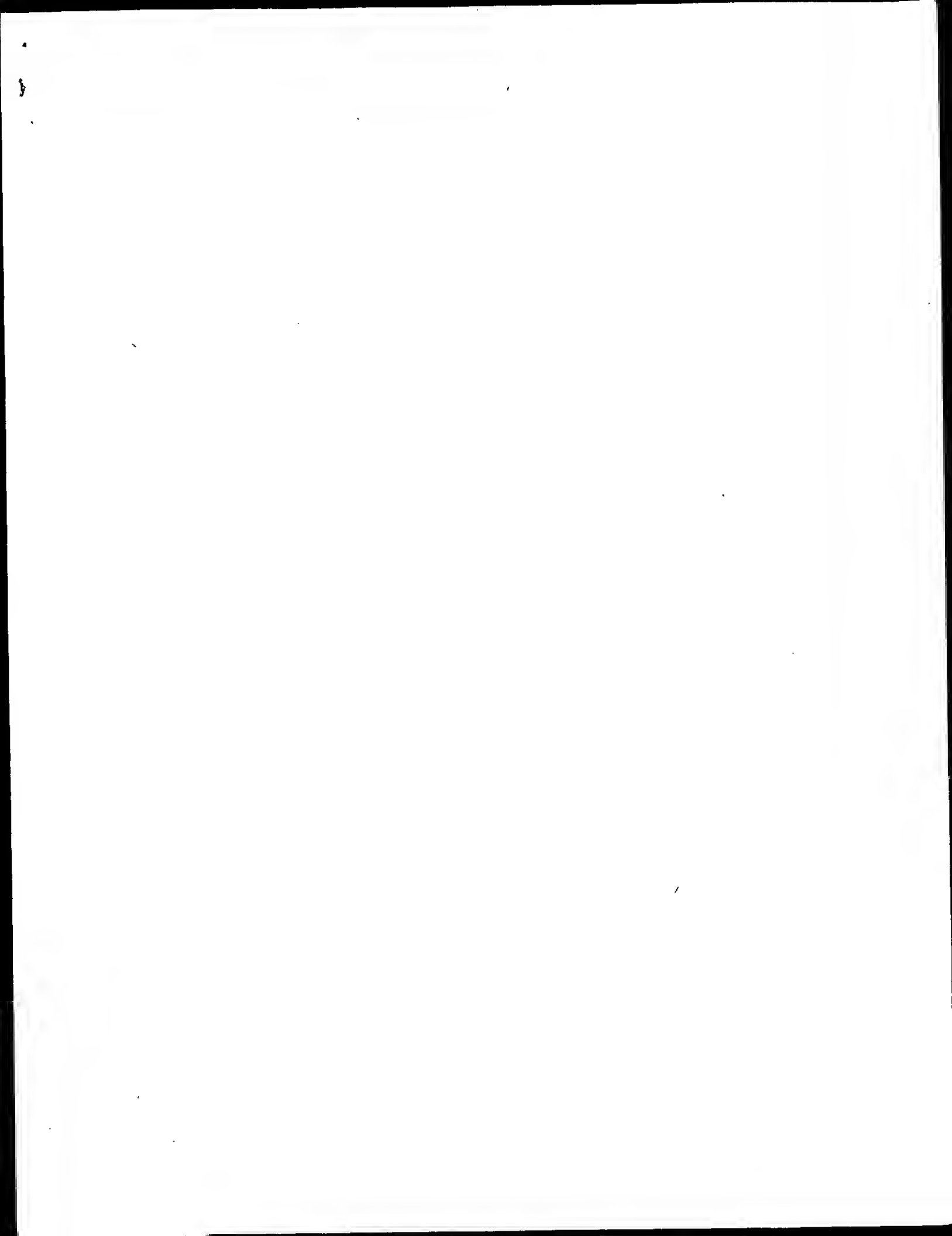
RESULT 15
US-09-105-678A-34
; Sequence 34, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,678A
; FILING DATE: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440

```

: INFORMATION FOR SEQ ID NO: 34:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 20 amino acids
 : TYPE: amino acid
 : STRANDEDNESS:
 : TOPOLOGY: linear
 : MOLECULE TYPE: peptide
 : US-09-105-678A-34

Query Match 66.9%; Score 115; DB 3; Length 20;
 Best Local Similarity 100.0%; Pred. No. 2.1e-11;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 12 TPDINPAMYAGRGIRPYGRF 31
 ||||||||||||||||
 Db 1 TPDINPAMYAGRGIRPYGRF 20

Search completed: April 17, 2001, 15:39:47
 Job time: 316 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 17, 2001, 15:32:53 ; Search time 116.94 Seconds
(without alignments)
15.154 Million cell updates/sec

Title: us-09-446-543a-5

Perfect score: 172
Sequence: 1 SRHGHSMETRPDIPNAPYAGRGIRVGRF 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A.GeneSeq_0401.*
1: /SIDSI/gcgdata/geneSeq/geneSeq/AA1980.DAT.*
2: /SIDSI/gcgdata/geneSeq/geneSeq/AA1981.DAT.*
3: /SIDSI/gcgdata/geneSeq/geneSeq/AA1982.DAT.*
4: /SIDSI/gcgdata/geneSeq/geneSeq/AA1983.DAT.*
5: /SIDSI/gcgdata/geneSeq/geneSeq/AA1984.DAT.*
6: /SIDSI/gcgdata/geneSeq/geneSeq/AA1985.DAT.*
7: /SIDSI/gcgdata/geneSeq/geneSeq/AA1986.DAT.*
8: /SIDSI/gcgdata/geneSeq/geneSeq/AA1987.DAT.*
9: /SIDSI/gcgdata/geneSeq/geneSeq/AA1988.DAT.*
10: /SIDSI/gcgdata/geneSeq/geneSeq/AA1989.DAT.*
11: /SIDSI/gcgdata/geneSeq/geneSeq/AA1990.DAT.*
12: /SIDSI/gcgdata/geneSeq/geneSeq/AA1991.DAT.*
13: /SIDSI/gcgdata/geneSeq/geneSeq/AA1992.DAT.*
14: /SIDSI/gcgdata/geneSeq/geneSeq/AA1993.DAT.*
15: /SIDSI/gcgdata/geneSeq/geneSeq/AA1994.DAT.*
16: /SIDSI/gcgdata/geneSeq/geneSeq/AA1995.DAT.*
17: /SIDSI/gcgdata/geneSeq/geneSeq/AA1996.DAT.*
18: /SIDSI/gcgdata/geneSeq/geneSeq/AA1997.DAT.*
19: /SIDSI/gcgdata/geneSeq/geneSeq/AA1998.DAT.*
20: /SIDSI/gcgdata/geneSeq/geneSeq/AA1999.DAT.*
21: /SIDSI/gcgdata/geneSeq/geneSeq/AA2000.DAT.*
22: /SIDSI/gcgdata/geneSeq/geneSeq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	172	100.0	31	W31371	Bovine G protein-c
2	172	100.0	31	W97218	Bovine pituitary-d
3	172	100.0	31	W97213	Bovine 19P2 ligand
4	172	100.0	31	W95188	Bovine pituitary-d
5	172	100.0	31	B10347	Bovine oxytocin se
6	172	100.0	31	V49290	19P2 ligand peptid
7	172	100.0	31	V49298	19P2 ligand peptid
8	172	100.0	32	W31372	Bovine G protein-c
9	172	100.0	32	W95189	Bovine pituitary-d
10	172	100.0	32	B10348	Bovine oxytocin se
11	172	100.0	33	W31373	Bovine G protein-c

12	172	100.0	33	20	W95190
13	172	100.0	33	21	B10349
14	172	100.0	33	21	V49297
15	172	100.0	98	18	W31382
16	172	100.0	98	18	W31386
17	172	100.0	98	20	W97224
18	172	100.0	98	20	W97217
19	172	100.0	98	20	W95187
20	172	100.0	98	21	B10346
21	172	100.0	98	21	B10353
22	172	100.0	98	21	W31384
23	163	94.8	31	18	W31384
24	163	94.8	31	20	W97233
25	163	94.8	31	20	W97614
26	163	94.8	31	20	W95173
27	163	94.8	31	20	W95174
28	163	94.8	31	21	B10355
29	163	94.8	31	21	V87504
30	163	94.8	31	21	V49292
31	163	94.8	32	18	W31385
32	163	94.8	32	21	B10356
33	163	94.8	33	18	W31386
34	163	94.8	33	21	B10357
35	163	94.8	82	20	W95172
36	163	94.8	83	18	W31383
37	163	94.8	83	20	W97225
38	163	94.8	83	21	B10354
39	161	93.6	29	18	W31359
40	161	93.6	29	20	W95184
41	158	91.9	31	18	W31391
42	158	91.9	31	20	W97235
43	158	91.9	31	20	W97615
44	158	91.9	31	21	B10362
45	158	91.9	32	18	V49291

ALIGNMENTS

RESULT	ID	W31371	standard; Peptide: 31 AA.
XX	AC	W31371;	
XX	DE	06-APR-1998 (first entry)	
XX	DE	Bovine G protein-coupled receptor ligand peptide fragment 1.	
XX	KW	G protein-coupled receptor; ligand binding; pharmacological; modulator; pituitary; central nervous system; pancreas; prophylactic; therapeutic agent.	
XX	KW		
XX	OS	Bos taurus.	
XX	PN	W09724436-A2.	
XX	PD	10-JUL-1997.	
XX	PF	26-DEC-1996; 96MO-JR03821.	
XX	PR	18-SEP-1996; 96UP-0246573.	
XX	PR	28-DEC-1995; 95UP-034371.	
XX	PR	15-MAR-1996; 96UP-0059419.	
XX	PR	12-AUG-1996; 96UP-0211805.	
XX	PA	(TAKE) TAKEDA CHEM IND LTD.	
XX	PI	Fuji R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;	
XX	PI	Kawamata Y, Kitada C;	
XX	DR	WPI; 1997-363672/33.	
XX	DR	N-PSDB; V02394.	

Bovine pituitary-d
Bovine oxytocin se
19P2 ligand peptid
Bovine genome deri
Bovine G protein-c
Bovine genome-driv
Bovine pituitary-d
Bovine genome-deri
Bovine oxytocin se
Bovine oxytocin se
Rat type G protein
Rat type ligand po
Rat 19P2 ligand.
Murine pituitary-d
Murine pituitary-d
Rat oxytocin secre
Rat prolactin-rela
19P2 ligand peptid
Rat type G protein
Rat oxytocin secre
Rat type G protein
Rat oxytocin secre
Murine pituitary-d
Rat type G protein
Rat type ligand po
Rat oxytocin secre
Bovine G protein-c
Bovine pituitary-d
Human type G prote
Human type ligand
Human 19P2 ligand
Human oxytocin sec
19P2 ligand peptid
Human type G prote

XX Ligand peptide for G protein-coupled receptor - acts by modulating
 PT function in the central nervous system, pancreas and pituitary gland
 PS Claim 2: Page 160, 258pp: English.

XX This sequence represents a peptide fragment of a novel bovine pituitary
 CC derived ligand corresponding to amino acid residues 23 to 53 of the
 CC sequence in W97218 and is used in an assay to monitor ligand binding to
 CC the G protein-coupled receptor protein. Pharmaceutical compositions
 CC containing this ligand may be used as a pituitary function modulator, a
 CC central nervous system modulator or a pancreatic function modulator. This
 CC ligand could have specific applications as a prophylactic or therapeutic
 CC agent for dementia, depression, hyperkinetic syndrome, disturbance of
 CC consciousness, anxiety syndrome, schizophrenia, trauma, growth hormone
 CC secretory disease, hyper- and polyphagia, hypercholesterolaemia,
 CC hyperglycaemia, hyperlipidaemia, hyperprolactinaemia, diabetes,
 CC cancer, pancreatitis, renal disease, Turner's syndrome, neurosis,
 CC rheumatoid arthritis, spinal injury, transient brain ischaemia,
 CC amyotrophic lateral sclerosis, acute myocardial infarction,
 CC spherothelial degeneration, bone fracture, trauma, atopic dermatitis,
 CC osteoporosis, asthma, epilepsy, infertility and/or oligospermia. Assays
 CC can also be developed to screen compounds which are capable of altering
 CC the binding activity of the ligand thus affecting activation of the G
 CC protein-coupled receptor protein.

XX Sequence 31 AA;

Query Match 100.0%, Score 172, DB 18, Length 31,
 Best Local Similarity 100.0%; Pred. No. 3, 9e-19;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQSMETRPDINPAMYAGRGIRPVGRF 31
 ID W97218
 DB 1 srahqsmetirpdpinpawjagrgirpvgrf 31

RESULT 2
 ID W97218 standard; peptide; 31 AA.
 AC W97218;

DT 06-MAY-1999 (first entry)

DE Bovine pituitary-derived ligand polypeptide fragment.

XX Bovine pituitary-derived ligand; modulation, prolactin secretion;
 KW G protein-coupled receptor; GPCR; hypocoarctanism; gonocyst cacogenesis;
 KW menopausal syndrome; euthyroid; hypometabolism; lactation;
 KW pituitary adenomatosis; brain tumour; emmenopathy; autoimmune disease;
 KW prolactinoma; infertility; impotence; amenorrhea; galactorrhea;
 KW acromegaly; Chari-Frommel syndrome; Argonz-del Castillo syndrome;
 KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;
 KW contraceptive; placental function; chorioncarcinoma; hydatid mole;
 KW interruption mole; abortion; unfertilized fetus; abnormal saccharometabolism;
 KW abnormal lipidmetabolism; oxytocia.

XX Bos sp.

XX WO9856962-A1.

XX 30-DEC-1998.

PF 22-JUN-1998, 98WO-JU02765.

XX 23-JUN-1997, 97JP-0165437.

XX (TAKE) TAKEDA CHEM IND LTD.

PI Fujii R, Hinuma S, Kawamata Y, Matsumoto H,
 XX

DR WPI, 1999-105614/09.

XX Use of G protein-coupled receptor ligands - for modulating prolactin
 PT secretion or placental function, e.g. for treating menopausal
 PT syndrome, tumours, autoimmune disease or abnormal pregnancy

PS Claim 3, Page 135, 241pp: English.

XX The present sequence represents a bovine pituitary-derived ligand
 CC fragment. It is used in the course of the invention. The specification
 CC describes an agent for modulating prolactin secretion which comprises a
 CC ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)
 CC protein. The agents for promoting prolactin secretion can be used for
 CC treating or preventing hypocoarctanism, gonocyst cacogenesis, menopausal
 CC syndrome, euthyroid or hypometabolism. They can be used for promoting
 CC lactation in a domestic mammal and as an aphrodisiac. The agents for
 CC inhibiting prolactin secretion can be used for treating or preventing
 CC pituitary adenomatosis, brain tumour, emmenopathy, autoimmune disease,
 CC prolactinoma, infertility, impotence, amenorrhea, galactorrhea,
 CC acromegaly, Chari-Frommel syndrome, Argonz-del Castillo syndrome,
 CC Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dyszoospermia.
 CC The inhibitory agents can also be used as contraceptives. The agents for
 CC modulating placental function can be used for treating or preventing
 CC chorioncarcinoma, hydatid mole, interruption mole, abortion, unfertilized fetus,
 CC abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.

XX Sequence 31 AA;

Query Match 100.0%, Score 172, DB 20, Length 31,
 Best Local Similarity 100.0%; Pred. No. 3, 9e-19;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHQSMETRPDINPAMYAGRGIRPVGRF 31
 ID W87613
 DB 1 srahqsmetirpdpinpawjagrgirpvgrf 31

RESULT 3
 ID W87613 standard; peptide; 31 AA.
 AC W87613;

DT 29-MAR-1999 (first entry)

DE Bovine 19P2 ligand.

XX 19P2 ligand; G protein coupled receptor; pituitary;
 KW prolactin releasing peptide; cattle; dementia; breast cancer;
 KW therapy.

XX Bos taurus.

PN EP887417-A2.

PD 30-DEC-1998.

PF 25-JUN-1998, 98EP-0111725.

XX 27-JUN-1997, 97JP-0172118.

XX (TAKE) TAKEDA CHEM IND LTD.

PI Moriya T, Nishimura O, Suenaga M, Tanaka Y;

XX WPI, 1999-047884/05.

PT Producing a 19P2 pituitary G protein receptor ligand - by cleavage
 PT of a fusion protein, useful for preventing and treating dementia,
 PT breast cancer, renal failure and autoimmune disease

PS Claim 5, Page 34, 56pp: English.

XX This is the amino acid sequence of the bovine pituitary G
 CC protein-coupled receptor ligand 19P2L. A method suitable for
 CC commercial high-level production of 19P2L comprises expressing
 CC the ligand in host cells as a recombinant fusion protein e.g. with
 CC human basic fibroblast growth factor (see V83792-93) that has
 CC been modified to include an N-terminal cysteine residue. The
 CC ligand is released from the fusion by cyanation followed by
 CC ammonolysis. 19P2L has prolactin secretion-stimulating and (at
 CC high doses) prolactin secretion-inhibiting properties. It can be
 CC used in the treatment and prevention of various diseases including:
 CC senile dementia, cerebrovascular dementia, and dementia associated
 CC with: neurological disorders (e.g. Alzheimer's disease, Parkinson's
 CC disease, Pick's disease, Huntington's disease), infectious diseases
 CC (e.g. Creutzfeldt-Jakob's), endocrine or metabolic disease or
 CC toxicosis (e.g. hypothyroidism, vitamin B12 deficiency, alcoholism,
 CC intoxication by drugs, metal and organic compounds), tumorigenic
 CC diseases (e.g. brain tumour), traumatic diseases (e.g. chronic
 CC subarachnoid hemorrhage, and other types of dementia, depression,
 CC hyperactive child syndrome (microencephalopathy) and disturbance of
 CC consciousness. It is also useful for prevention and treatment of
 CC diseases associated with prolactin hypo and hypersecretion
 CC respectively, including: hyperprolactinemia, pituitary adenoma,
 CC breast cancer, infertility, impotence and autoimmune disease
 CC (hypersecretion disorders), and seminal vesicle hypoplasia,
 CC osteoporosis, menopausal syndrome and renal failure (hyposecretion
 CC disorders). The 19P2 polypeptide/amide is also useful as a test
 CC reagent for study of the prolactin secretory function or as a
 CC lactagogue in mammalian farm animals.

SQ Sequence 31 AA:

Query Match 100.0%; Score 172; DB 20; Length 31;
 Best Local Similarity 100.0%; Pred. No. 3.9e-19;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRAHSHMEIRTPDINPAWYAGRGIRPVGRF 31
 DB 1 strahshmeirtpdinpawyagrgirpvgrf 31

RESULT 4

ID W95188 standard; peptide; 31 AA.

AC W95188;

DT 10-MAR-1999 (first entry)

DE Bovine pituitary-derived ligand polypeptide fragment.

XX Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;
 KW GPR10; modulator; pituitary; central nervous system; pancreas;
 KW tissue; screen; therapeutic; binding; senile dementia; ligand; murine;
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;
 KW Creutzfeldt-Jakob disease; poisoning; schizophrenia; growth hormone;
 KW secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;
 KW gene therapy; transgenic animal; bovine.

XX Bos sp.

OS WO9849295-A1.

PN 05-NOV-1998.

PD 27-APR-1998; 98WO-JP01923.

PR 28-APR-1997; 97JP-0109974.

PA (TAKE) TAKEDA CHEM IND LTD.

PI Fukusumi S, Hinuma S;

XX WPL; 1999-009423/01.
 DR New polypeptide ligand for orphan G protein coupled receptors - used
 XX for treating disorders of central nervous system, pituitary and
 PT pancreas, and for drug screening
 PS Example 19; Page 150; 206pp; English.

XX The invention relates to a murine pituitary-derived ligand polypeptide
 CC which is a ligand for the G-protein coupled orphan receptor designated
 CC GPR10 (human) or UHR-1 (rat). Cells transformed with a vector containing
 CC the ligand polypeptide encoding DNA are used to produce a recombinant
 CC ligand polypeptide. The ligand polypeptide, and its fragments, modulate
 CC function of the pituitary, central nervous system, pancreas and other
 CC tissues and can be used to screen for agents that modulate binding of the
 CC polypeptide to the receptor; to quantify the amount of receptor in a
 CC sample and to raise antibodies. They may also be used therapeutically,
 CC e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's
 CC diseases; Creutzfeldt-Jakob disease; poisoning by heavy metals or drugs;
 CC diabetes; schizophrenia; disorders of growth hormone secretion; cancer;
 CC rheumatoid arthritis, epilepsy and many others, also to improve post-
 CC operative nutritional status and as vasopressor. Transgenic animals
 CC carrying the ligand polypeptide encoding DNA or its mutagen are used to
 CC study the function of the polypeptide-expressing genes, as models of
 CC disease, for drug screening and as source of cell lines. The ligand
 CC polypeptide DNA is used as a source of probes and primers; to identify
 CC related sequences; in receptor-binding assays; for production of Ab and
 CC antisera; in drug development; for gene therapy and to develop transgenic
 CC animals. The present sequence represents a bovine genome-derived ligand
 CC polypeptide fragment which is similar to the murine ligand polypeptide.

SQ Sequence 31 AA:

Query Match 100.0%; Score 172; DB 20; Length 31;
 Best Local Similarity 100.0%; Pred. No. 3.9e-19;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRAHSHMEIRTPDINPAWYAGRGIRPVGRF 31
 DB 1 strahshmeirtpdinpawyagrgirpvgrf 31

RESULT 5

ID B10347 standard; peptide; 31 AA.

AC B10347;

DT 24-NOV-2000 (first entry)

DE Bovine oxytocin secretion promoting peptide SEQ ID NO: 3.

XX Bovine; oxytocin secretion promoter; G protein-coupled receptor protein;
 KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;
 KW caesarean section; artificial fertilization; galactostasis; goat; pig;
 KW veterinary medicine; milk production.

XX Bos taurus.

OS WO200038704-A1.

PN 06-JUL-2000.

PD 22-DEC-1999; 99WO-JP07199.

PR 25-DEC-1998; 98JP-0369585.

PA (TAKE) TAKEDA CHEM IND LTD.

PI Matsumoto H, Kitada C, Hinuma S;

DR	WPI; 2000-452298/39.
XX	
PT	Physiologically-active polypeptide recognized as ligand by G
PT	protein-coupled receptor protein, for promoting secretion of oxytocin,
PT	as drugs for diseases relating to oxytocin secretion and in veterinary
PT	medicine -
XX	
PS	Claim 3; Page 50; 72pp; Japanese.
CC	
CC	This invention describes a novel oxytocin secretion-regulating agent
CC	which contains a ligand peptide or its salt for the G protein-coupled
CC	receptor protein. It is useful in the form of drugs for ameliorating,
CC	preventing and treating diseases relating to oxytocin secretion e.g.,
CC	weak pains and atonic bleeding, before and after expulsion of placenta,
CC	uterine recovery failure, caesarean section, stoppage of artificial
CC	fertilization or galactostasis and is also applicable in veterinary
CC	medicine for promoting milk production in cow, goat and pig. This
CC	sequence represents a bovine peptide which acts as an oxytocin secretion
CC	promoter.
XX	
SQ	Sequence 31 AA:
	Query Match 100.0%; Score 172; DB 21; Length 31;
	Best Local Similarity 100.0%; Pred. No. 3,9e-19;
	Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	1 SRAHQHMEIRTPDINPAMVAGRGIRPGGRF 31
Db	1 strahqhmeirtpdinpawagrgirpgrf 31
RESULT 6	
ID Y49290	Y49290 standard; peptide; 31 AA.
AC XX	
XX Y49290;	
DT 22-FEB-2000	(first entry)
DE 19P2 ligand peptide fragment.	
FW Monoclonal antibody; 19P2 ligand; diagnosis; prolactin secretion;	
KW pluriatary; regulatory mechanism; central nervous system; pancreatic.	
OS Bos sp.	
XX Key Location/Qualifiers	
FH Modified-site 31	
FT /note= "C-terminal amide"	
FN W09960112-A1.	
PD 25-NOV-1999.	
XX 20-MAY-1999;	99WO-JP02650.
XX 21-MAY-1998;	98JP-0140293.
PA (TAKE) TAKEDA CHEM IND LTD.	
PI Matsumoto H, Kitada C, Hinuma S;	
DR WPI; 2000-039381/03.	
XX	
PT New monoclonal antibodies, useful in diagnosis, as drugs and in	
PT studying diseases related to ligand abnormality -	
PS Disclosure; Page 26; 73pp; Japanese.	
CC The invention provides a monoclonal antibody which has a specific	
CC reaction with the part peptide of the C-terminal of 19P2 ligand or its	
CC derivative. The antibodies can be used in diagnosis or to treat or	

```

CC    prenu-diseases associated with abnormality in the pituitary function
CC    regulatory mechanism (e.g. promotion of prolactin secretion), central
CC    nervous regulatory mechanism, and pancreatic function regulatory
CC    mechanism. The antibody-based immunoassay can also be applied in
CC    clarifying the physiological functions of the ligand and its derivative.
CC    Sequences Y49290-302 represent peptide fragments of the 19p2 ligand.
CC    XX
SQ    Sequence      31 AA;

Query Match      100.0%; Score 172; DB 21; Length 31;
Best Local Similarity 100.0%; Pred. No. 3.9e-19;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY    1 SRAHQSMSEIRPTDINPAWYAGRCIRPVGRF 31
      |||
      1 srahqsmseirptdinpawyagrcirpvgrf 31

RESULT 7
ID    Y49298 standard; peptide; 31 AA.
XX
AC    Y49298;
XX
DT    22-FEB-2000 (first entry)
XX
DE    19p2 ligand peptide fragment.
XX
KW    Monoclonal antibody; 19p2 ligand; diagnosis; prolactin secretion;
XX    pituitary; regulatory mechanism; central nervous system; pancreatic.
XX
OS    Homo sapiens.
XX
FH    Key Location/Qualifiers
FT    Modified-site 31
FT    /note="C-terminal amide"
XX
PN    WO9960112-A1.
XX
PD    25-NOV-1999.
XX
PF    20-MAY-1999; 99MO-JP02650.
XX
PR    21-MAY-1998; 98JP-0140293.
XX
PT    (TAKE ) TAKEDA CHEM IND LTD.
XX
PI    Matsumoto H, Kitada C, Hinuma S;
XX    WPI; 2000-039381/03.
XX
DR    New monoclonal antibodies, useful in diagnosis, as drugs and in
XX    studying diseases related to ligand abnormality .
XX
PS    Disclosure; Page 27; 73pp; Japanese.
XX
CC    The invention provides a monoclonal antibody which has a specific
CC    reaction with the part peptide of the C-terminal of 19p2 ligand or its
CC    derivative. The antibodies can be used in diagnosis or to treat or
CC    prevent diseases associated with abnormality in the pituitary function
CC    regulatory mechanism (e.g. promotion of prolactin secretion), central
CC    nervous regulatory mechanism, and pancreatic function regulatory
CC    mechanism. The antibody-based immunoassay can also be applied in
CC    clarifying the physiological functions of the ligand and its derivative.
CC    Sequences Y49290-302 represent peptide fragments of the 19p2 ligand.
CC    XX
SQ    Sequence      31 AA;

Query Match      100.0%; Score 172; DB 21; Length 31;
Best Local Similarity 100.0%; Pred. No. 3.9e-19;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Oy 1 SRAHSHMEIRTPDINPAWYAGRGIRPVGRF 31
 |||||
 Db 1 srahshmeirtpdinpawyagrgirpvgrf 31

RESULT 8
 ID W31372 standard; peptide; 32 AA.
 W31372:

AC W31372:
 DT 06-APR-1998 (first entry)

DE Bovine G protein-coupled receptor ligand peptide fragment 2.

KW G protein-coupled receptor; ligand binding; pharmaceutical;
 KW modulator; pituitary; central nervous system; pancreas; prophylactic;
 KW therapeutic agent.

OS Bos taurus.

PN WO9724436-A2.

PD 10-JUL-1997.

PF 26-DEC-1996; 96WO-JP03821.

PR 18-SEP-1996; 96JP-0246573.

PR 28-DEC-1995; 95JP-0343371.

PR 15-MAR-1996; 96JP-0059419.

PR 12-AUG-1996; 96JP-0211805.

PA (TAKE) TAKEDA CHEM IND LTD.

PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M,
 PI Kawamata Y, Kitada C;

DR WPI, 1997-363672/33.

DR N-PSDB; V02395.

PT Ligand peptide for G protein-coupled receptor - acts by modulating
 PT function in the central nervous system, pancreas and pituitary gland

PS Claim 2; Page 160-161, 258pp; English.

XX This sequence represents a peptide fragment of a novel bovine pituitary
 CC derived ligand corresponding to amino acid residues 23 to 54 of the
 CC sequence in W31368 and is used in an assay to monitor ligand binding to
 CC the G protein-coupled receptor protein. Pharmaceutical compositions
 CC containing this ligand may be used as a pancreatic function modulator, a
 CC central nervous system modulator or a pancreatic function modulator. This
 CC agent could have specific applications as a prophylactic or therapeutic
 CC agent for dementia, depression, hyperkinetic syndrome, disturbance of
 CC consciousness, anxiety syndrome, schizophrenia, trauma, growth hormone
 CC secretory disease, hyper- and polyphagia, hypercholesterolaemia,
 CC hyperglycaemia, hyperlipidaemia, hyperprolactinaemia, diabetes,
 CC cancer, pancreatitis, renal disease, Turner's syndrome, neurosis,
 CC rheumatoid arthritis, spinal injury, transient brain ischaemia,
 CC amyloidotic lateral sclerosis, acute myocardial infarction,
 CC aplodysplasia, degeneration, bone fracture, trauma, atopic dermatitis,
 CC osteoporosis, asthma, epilepsy, infertility and/or oligosaccharia. Assays
 CC can also be developed to screen compounds which are capable of altering
 CC the binding activity of the ligand thus affecting activation of the G
 CC protein-coupled receptor protein.

XX Sequence 32 AA;

Query Match 100.0%; Score 172; DB 18; Length 32;
 Best Local Similarity 100.0%; Pred. No. 4, 1e-19;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 SRAHSHMEIRTPDINPAWYAGRGIRPVGRF 31
 |||||
 Db 1 srahshmeirtpdinpawyagrgirpvgrf 31

RESULT 9
 ID W95189 standard; peptide; 32 AA.
 W95189:

AC W95189:
 DT 10-MAR-1999 (first entry)

DE Bovine pituitary-derived ligand polypeptide fragment.

KW Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;
 KW GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas;
 KW tissue; screen; therapeutic; binding; senile dementia; ligand; murine;
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;
 KW Creutzfeldt-Jakob disease; poisoning; schizophrenia; growth hormone;
 KW secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;
 KW gene therapy; transgenic animal; bovine.

OS Bos sp.

PN WO9849295-A1.

PD 05-NOV-1998.

PF 27-APR-1998; 98WO-JP01923.

PR 28-APR-1997; 97JP-0109974.

PA (TAKE) TAKEDA CHEM IND LTD.

PI Fukusumi S, Hinuma S;

DR WPI, 1999-009423/01.

PT New polypeptide ligand for orphan G protein coupled receptors - used
 PT for treating disorders of central nervous system, pituitary and
 PT pancreas, and for drug screening

PS Example 19; Page 150; 206pp; English.

XX The invention relates to a murine pituitary-derived ligand polypeptide
 CC which is a ligand for the G-protein coupled orphan receptor designated
 CC GPR10 (human) or UHR-1 (rat). Cells transformed with a vector containing
 CC the ligand polypeptide encoding DNA are used to produce a recombinant
 CC ligand polypeptide. The ligand polypeptide, and its fragments, modulate
 CC function of the pituitary, central nervous system, pancreas and other
 CC tissues and can be used to screen for agents that modulate binding of the
 CC polypeptide to the receptor; to quantify the amount of receptor in a
 CC sample and to raise antibodies. They may also be used therapeutically,
 CC e.g. to treat senile dementia, Alzheimer's, Parkinson's or Huntington's
 CC diseases; Creutzfeldt-Jakob disease; poisoning by heavy metals or drugs;
 CC diabetes; schizophrenia; disorders of growth hormone secretion; cancer;
 CC rheumatoid arthritis; epilepsy and many others, also to improve post-
 CC operative nutritional status and as vasopressor. Transgenic animals
 CC carrying the ligand polypeptide encoding DNA or its mutant are used to
 CC study the function of the polypeptide-expressing genes, as models of
 CC disease, for drug screening and as source of cell lines. The ligand
 CC polypeptide DNA is used as a source of probes and primers; to identify
 CC related sequences; in receptor-binding assays; for production of Ab and
 CC antisera; in drug development; for gene therapy and to develop transgenic
 CC animals. The present sequence represents a bovine genome-derived ligand
 CC polypeptide fragment which is similar to the murine ligand-polypeptide.

XX Sequence 32 AA;

Query Match 100.0%; Score 172; DB 20; Length 32;
 Best Local Similarity 100.0%; Pred. No. 4, 1e-19;

Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHSHMEIRTPDINPAMYAGRGIRPVGRF 31
 |||
 DB 1 srahshmeirtpdinpawagrgirpvgrf 31

RESULT 10
 ID B10348 standard; peptide; 32 AA.
 XX
 AC B10348;
 XX
 DT 24-NOV-2000 (first entry)
 XX
 DE Bovine oxytocin secretion promoting peptide SEQ ID NO: 4.
 XX
 KW Bovine; oxytocin secretion promoter; G protein-coupled receptor protein;
 KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;
 KW caesarean section; artificial fertilization; galactostasis; goat; pig;
 KW veterinary medicine; milk production.
 XX
 OS Bos taurus.
 XX
 PA MO200038704-A1.
 XX
 PD 06-JUL-2000.
 XX
 PF 22-DEC-1999; 99WO-JP07199.
 XX
 PR 25-DEC-1998; 98JP-0369585.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX
 PI Matsumoto H, Kitada C, Hinuma S;
 XX
 DR WPI; 2000-452298/39.
 XX
 PT Physiologically-active polypeptide recognized as ligand by G
 PT protein-coupled receptor protein, for promoting secretion of oxytocin,
 PT as drugs for diseases relating to oxytocin secretion and in veterinary
 PT medicine -
 XX
 PS Disclosure; Page 51; 72pp; Japanese.
 XX
 CC This invention describes a novel oxytocin secretion-regulating agent
 CC which contains a ligand peptide or its salt for the G protein-coupled
 CC receptor protein. It is useful in the form of drugs for ameliorating,
 CC preventing and treating diseases relating to oxytocin secretion e.g.,
 CC weak pains and atonic bleeding, before and after expulsion of placenta,
 CC uterine recovery failure, caesarean section, stoppage of artificial
 CC fertilization or galactostasis and is also applicable in veterinary
 CC medicine for promoting milk production in cow, goat and pig. This
 CC sequence represents a bovine peptide which acts as an oxytocin secretion
 CC promoter.
 XX
 SQ Sequence 32 AA;

Query Match 100.0%; Score 172; DB 21; Length 32;
 Best Local Similarity 100.0%; Pred. No. 4, 1e-19;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHSHMEIRTPDINPAMYAGRGIRPVGRF 31
 |||
 DB 1 srahshmeirtpdinpawagrgirpvgrf 31

RESULT 11
 ID W31373 standard; Peptide; 33 AA.
 XX
 AC W31373;

XX
 DT 06-APR-1998 (first entry)
 XX
 DE Bovine G protein-coupled receptor ligand peptide fragment 3.
 XX
 KW G protein-coupled receptor; ligand binding; pharmaceutical;
 KW modulator; pituitary; central nervous system; pancreas; prophylactic;
 KW therapeutic agent.
 XX
 OS Bos taurus.
 XX
 PA MO9724436-A2.
 XX
 PD 10-JUL-1997.
 XX
 PF 26-DEC-1996; 96WO-JP03821.
 XX
 PR 18-SEP-1996; 96JP-0246573.
 PR 28-DEC-1995; 95JP-034371.
 PR 15-MAR-1996; 96JP-0059419.
 PR 12-AUG-1996; 96JP-0211805.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX
 PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;
 PI Kawamata Y, Kitada C;
 XX
 DR WPI; 1997-363672/33.
 DR N-PSDB; V02396.
 XX
 PT Ligand peptide for G protein-coupled receptor - acts by modulating
 PT function in the central nervous system, pancreas and pituitary gland .
 XX
 PS Claim 2; Page 161; 258pp; English.
 XX
 CC This sequence represents a peptide fragment of a novel bovine pituitary
 CC derived ligand corresponding to amino acid residues 23 to 55 of the
 CC sequence in W31368 and is used in an assay to monitor ligand binding to
 CC the G protein-coupled receptor protein. Pharmaceutical compositions
 CC containing this ligand may be used as a pituitary function modulator, a
 CC central nervous system modulator or a pancreatic function modulator. This
 CC ligand could have specific applications as a prophylactic or therapeutic
 CC agent for dementia, depression, hyperkinetic syndrome, disturbance of
 CC consciousness, anxiety syndrome, schizophrenia, trauma, growth hormone
 CC secretory disease, hyper- and polyphagia, hypercholesterolaemia,
 CC hyperglycidaemia, hyperlipidaemia, hyperprolactinaemia, diabetes,
 CC cancer, pancreatitis, renal disease, Turner's syndrome, neurosis,
 CC rheumatoid arthritis, spinal injury, transient brain ischaemia,
 CC amyotrophic lateral sclerosis, acute myocardial infarction,
 CC spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis,
 CC osteoporosis, asthma, epilepsy, infertility and/or oligosaccharia. Assays
 CC can also be developed to screen compounds which are capable of altering
 CC the binding activity of the ligand thus affecting activation of the G
 CC protein-coupled receptor protein.
 XX
 SQ Sequence 33 AA;

Query Match 100.0%; Score 172; DB 18; Length 33;
 Best Local Similarity 100.0%; Pred. No. 4, 2e-19;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHSHMEIRTPDINPAMYAGRGIRPVGRF 31
 |||
 DB 1 srahshmeirtpdinpawagrgirpvgrf 31

RESULT 12
 ID W95190 standard; peptide; 33 AA.
 XX
 AC W95190;

DT 10-MAR-1999 (first entry)
 XX
 DE Bovine pituitary-derived ligand polypeptide fragment.
 XX
 KW pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;
 KW GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas;
 KW tissue; screen; therapeutic; binding; senile dementia; ligand; murine;
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;
 KW Creutzfeldt-Jakob disease; poisoning; schizophrenia; growth hormone;
 KW secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;
 KW gene therapy; transgenic animal; bovine.
 XX
 OS Bos sp.
 XX
 XX MO9849295-A1.
 XX
 PD 05-NOV-1998.
 XX
 XX 27-APR-1998; 98MO-JP01923.
 XX
 PF 28-APR-1997; 97JP-0109974.
 XX
 PR (TAKE) TAKEDA CHEM IND LTD.
 XX
 PA Fukusumi S, Hinuma S;
 XX
 PI WPI; 1999-009423/01.
 XX
 DR New polypeptide ligand for orphan G protein coupled receptors - used
 XX
 PT for treating disorders of central nervous system, pituitary and
 XX
 PT pancreas, and for drug screening
 XX
 PS Example 19; Page 150; 206pp; English.
 XX
 CC The invention relates to a murine pituitary-derived ligand polypeptide
 CC which is a ligand for the G-protein coupled orphan receptor designated
 CC GPR10 (human) or UHR-1 (rat). Cells transformed with a vector containing
 CC the ligand polypeptide encoding DNA are used to produce a recombinant
 CC ligand polypeptide. The ligand polypeptide, and its fragments, modulate
 CC function of the pituitary, central nervous system, pancreas and other
 CC tissues and can be used to screen for agents that modulate binding of the
 CC polypeptide to the receptor; to quantify the amount of receptor in a
 CC sample and to raise antibodies. They may also be used therapeutically,
 CC e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's
 CC diseases; Creutzfeldt-Jakob disease; poisoning by heavy metals or drugs;
 CC diabetes; schizophrenia; disorders of growth hormone secretion; cancer;
 CC rheumatoid arthritis; epilepsy and many others, also to improve post-
 CC operative nutritional status and as vasopressor. Transgenic animals
 CC carrying the ligand polypeptide encoding DNA or its mutagen are used to
 CC study the function of the polypeptide-expressing genes, as models of
 CC disease, for drug screening and as source of cell lines. The ligand
 CC polypeptide DNA is used as a source of probes and primers to identify
 CC related sequences; in receptor-binding assays; for production of Ab and
 CC antisera; in drug development; for gene therapy and to develop transgenic
 CC animals. The present sequence represents a bovine genome-derived ligand
 CC polypeptide fragment which is similar to the murine ligand-polypeptide.
 XX
 SO Sequence 33 AA;
 XX
 Query Match 100.0%; Score 172; DB 20; Length 33;
 Best local Similarity 100.0%; Pred. No. 4.2e-19;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SRAHSHMEIRTPDINPMWYAGRGIRPVGRF 31
 DB 1 srahshmeirtpdinpmwyagr9irpvgrf 31
 RESULT 13
 ID B10349
 XX B10349 standard; peptide; 33 AA.

AC B10349;
 XX
 DE 24-NOV-2000 (first entry)
 XX
 DE Bovine oxytocin secretion promoting peptide SEQ ID NO: 5.
 XX
 KW Bovine; oxytocin secretion promoter; G protein-coupled receptor protein;
 KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;
 KW caesarean section; artificial fertilization; galactostasis; goat; pig;
 KW veterinary medicine; milk production.
 XX
 OS Bos taurus.
 XX
 XX MO200038704-A1.
 XX
 PD 06-JUL-2000.
 XX
 XX 22-DEC-1999; 99MO-JP07199.
 XX
 PF 25-DEC-1998; 98JP-0369585.
 XX
 PR (TAKE) TAKEDA CHEM IND LTD.
 XX
 PA Matsumoto H, Kitada C, Hinuma S;
 XX
 PI WPI; 2000-452298/39.
 XX
 DR Physiologically-active polypeptide recognized as ligand by G
 XX
 PT protein-coupled receptor protein, for promoting secretion of oxytocin,
 XX
 PT as drugs for diseases relating to oxytocin secretion and in veterinary
 XX
 PT medicine
 XX
 PS Disclosure; Page 51; 72pp; Japanese.
 XX
 CC This invention describes a novel oxytocin secretion-regulating agent
 CC which contains a ligand peptide or its salt for the G protein-coupled
 CC receptor protein. It is useful in the form of drugs for ameliorating,
 CC preventing and treating diseases relating to oxytocin secretion e.g.,
 CC weak pains and atonic bleeding, before and after expulsion of placenta,
 CC uterine recovery failure, caesarean section, stoppage of artificial
 CC fertilization or galactostasis and is also applicable in veterinary
 CC medicine for promoting milk production in cow, goat and pig. This
 CC sequence represents a bovine peptide which acts as an oxytocin secretion
 CC promoter.
 XX
 SO Sequence 33 AA;
 XX
 Query Match 100.0%; Score 172; DB 21; Length 33;
 Best local Similarity 100.0%; Pred. No. 4.2e-19;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SRAHSHMEIRTPDINPMWYAGRGIRPVGRF 31
 DB 1 srahshmeirtpdinpmwyagr9irpvgrf 31
 RESULT 14
 ID Y49297
 XX Y49297 standard; peptide; 33 AA.
 AC Y49297;
 XX
 DE 22-FEB-2000 (first entry)
 XX
 DE 19P2 ligand peptide fragment.
 XX
 KW Monoclonal antibody; 19P2 ligand; diagnosis; prolactin secretion;
 KW pituitary; regulatory mechanism; central nervous system; pancreatic.
 XX
 OS Bos sp.
 XX
 XX MO9960112-A1.

XX 25-NOV-1999.
 PD 20-MAY-1999; 99WO-JP02650.
 XX
 PF 21-MAY-1998; 98JP-0140293.
 XX
 PR (TAKE) TAKEDA CHEM IND LTD.
 PA
 PI Matsumoto H, Kitada C, Hinuma S;
 XX WPI; 2000-039381/03.
 DR
 XX
 PT New monoclonal antibodies, useful in diagnosis, as drugs and in
 PT studying diseases related to ligand abnormality -
 PS Disclosure; Page 27; 73pp; Japanese.
 XX
 CC The invention provides a monoclonal antibody which has a specific
 CC reaction with the part peptide of the C-terminal of 19p2 ligand or its
 CC derivative. The antibodies can be used in diagnosis or to treat or
 CC prevent diseases associated with abnormality in the pituitary function
 CC regulatory mechanism (e.g. promotion of prolactin secretion), central
 CC nervous regulatory mechanism, and pancreatic function regulatory
 CC mechanism. The antibody-based immunoassay can also be applied in
 CC clarifying the physiological functions of the ligand and its derivative.
 CC Sequences Y49290-302 represent peptide fragments of the 19p2 ligand.
 CC
 SQ Sequence 33 AA;

Query Match 100.0%; Score 172; DB 21; Length 33;
 Best Local Similarity 100.0%; Pred. No. 4.2e-19;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRAHGSMEIRTPDINPAMYAGRGIRPVGRF 31
 ||||||||||||||||||||||||||||||||
 DB 1 strahgsmeirtpdinpawyagrgrirpvgrf 31

RESULT 15
 W31382
 ID W31382 standard; Protein; 98 AA.
 XX
 AC W31382;
 XX
 DT 06-APR-1998 (first entry)
 XX
 DE Bovine genome derived G protein-coupled receptor ligand.
 XX
 KW G protein-coupled receptor; ligand binding; pharmaceutical;
 KW modulator; pituitary; central nervous system; pancreas; prophylactic;
 KW therapeutic agent.
 XX
 OS Bos taurus.
 OS
 PN WO9724436-A2.
 PD 10-JUL-1997.
 XX
 PF 26-DEC-1996; 96WO-JP03821.
 XX
 PR 18-SEP-1996; 96JP-0246573.
 PR 28-DEC-1995; 95JP-0343371.
 PR 15-MAR-1996; 96JP-0059419.
 PR 12-AUG-1996; 96JP-0211805.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX
 PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;
 PI Kawamata Y, Kitada C;
 XX
 DR WPI; 2000-039381/03.

XX Ligand peptide for G protein-coupled receptor - acts by modulating
 PT function in the central nervous system, pancreas and pituitary gland
 XX
 PS Claim 3; Page 177; 258pp; English.

This sequence represents a novel bovine genome-derived ligand polypeptide
 CC which is used in an assay to monitor ligand binding to the G
 CC protein-coupled receptor protein. Pharmaceutical compositions containing
 CC this ligand may be used as a pituitary function modulator, a central
 CC nervous system modulator or a pancreatic function modulator. This ligand
 CC could have specific applications as a prophylactic or therapeutic agent
 CC for dementia, depression, hyperkinetic syndrome, disturbance of
 CC consciousness, anxiety syndrome, schizophrenia, trauma, growth hormone
 CC secretory disease, hyper- and polyphagia, hypercholesterolemia,
 CC hyperglycemia, hyperlipidemia, hyperprolactinemia, diabetes,
 CC cancer, pancreatitis, renal disease, Turner's syndrome, neurosis,
 CC rheumatoid arthritis, spinal injury, transient brain ischaemia,
 CC amyotrophic lateral sclerosis, acute myocardial infarction,
 CC spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis,
 CC osteoporosis, asthma, epilepsy, infertility and/or oligospermia. Assays
 CC can also be developed to screen compounds which are capable of altering
 CC the binding activity of the ligand thus affecting activation of the G
 CC protein-coupled receptor protein.
 CC
 SQ Sequence 98 AA;

Query Match 100.0%; Score 172; DB 18; Length 98;
 Best Local Similarity 100.0%; Pred. No. 1.6e-18;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRAHGSMEIRTPDINPAMYAGRGIRPVGRF 31
 ||||||||||||||||||||||||||||||||
 DB 23 strahgsmeirtpdinpawyagrgrirpvgrf 53

Search completed: April 17, 2001, 15:38:41
 Job time: 348 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 17, 2001, 15:35:52 ; Search time 70.08 Seconds
(without alignments)
30.400 Million cell updates/sec

Title: US-09-446-543A-5

Perfect score: 172

Sequence: 1 SRAHQSMETRPDINPAWYAGCIRPVGR 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

PIR 67: **
1: pirl: **
2: pirl: **
3: pirl: **
4: pirl: **

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	32.0	1236	2 T50904	Mg protoporphyrin
2	54.5	31.7	503	2 A82193	Sun/nucleolar prot
3	53.5	31.1	176	2 S67150	hypothetical prote
4	53	30.8	376	2 C75580	adenine deaminase-
5	52	30.2	1292	2 T31462	probable magnesium
6	51.5	29.9	664	2 F83376	conserved hypotbet
7	51	29.7	294	2 T21075	hypothetical prote
8	51	29.7	430	1 B69009	conserved hypotbet
9	51	29.7	798	2 S11210	probable untr prote
10	51	29.7	962	2 H69157	exclunuclease ABC c
11	50.5	29.4	790	2 T47959	hypothetical prote
12	50	29.1	527	2 T33175	hypothetical prote
13	49	28.5	128	2 S76955	hypothetical prote
14	48.5	28.2	254	2 S76814	hypothetical prote
15	48.5	28.2	348	2 T47548	hypothetical prote
16	48.5	28.2	1882	2 S73484	hypothetical prote
17	48	27.9	220	2 C83292	probable giutathio
18	48	27.9	314	2 B70569	hypothetical prote
19	48	27.9	348	2 T21648	hypothetical prote
20	48	27.9	365	2 T39098	hypothetical prote
21	48	27.9	424	2 B38176	samb protein - Sal
22	48	27.9	772	2 T07958	protoporphyrin IX
23	48	27.9	960	2 A71315	exclunuclease ABC c
24	48	27.9	1331	2 S75000	protoporphyrin IX
25	48	27.9	1379	2 S37310	protoporphyrin IX
26	48	27.9	1380	2 S64721	protoporphyrin IX
27	48	27.9	1381	2 S71288	protoporphyrin IX
28	48	27.9	1382	2 T01789	protoporphyrin IX
29	48	27.9	1383	2 T07126	magnesium chelatase

30	47	27.3	406	2 I46535	11-beta-hydroxyste
31	47	27.3	455	2 D70885	probable aldC prot
32	47	27.3	785	2 F69099	sensory transducti
33	47	27.3	940	2 A82329	exclunuclease ABC,
34	47	27.3	972	2 A70619	exclunuclease ABC c
35	47	27.3	1084	2 T33759	hypothetical prote
36	46.5	27.0	345	2 D84012	N-acetylglutamate
37	46.5	27.0	531	2 T35226	nitrate reductase
38	46.5	27.0	770	1 A38230	inorganic pyrophos
39	46.5	27.0	957	2 A84089	hypothetical prote
40	46	26.7	333	2 H82852	hydroxybenzoate oc
41	46	26.7	342	2 B64395	malic acid transpo
42	46	26.7	347	2 H64371	malic acid transpo
43	46	26.7	510	2 A64865	ycgb protein - Esc
44	46	26.7	798	2 S29815	N-ras upstream pro
45	46	26.7	836	2 A54269	protein-glutamine

ALIGNMENTS

RESULT 1
T50904
Mg protoporphyrin methyl transferase [imported] - Rubrivivax gelatinosus
C:Species: Rubrivivax gelatinosus
C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 02-Sep-2000
C:Accession: T50904
R:Nagashima, K.V.; Igarashi, N.; Harada, J.; Nagashima, S.; Matsuura, K.; Shimada, K.
Submitted to the EMBL Data Library, November 1999
A:Description: Determination of Nucleotide Sequences of Rubrivivax gelatinosus Photos
A:Reference number: 225270
A:Accession: T50904
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1236 <NAC>
A:Cross-references: EMBL:AB034704; PDB:BA94057.1
A:Experimental source: strain T1144
C:Genetics:
A:Gene: bchH
C:Superfamily: Rhodobacter capsulatus magnesium-protoporphyrin O-methyltransferase

Query Match 32.0%; Score 55; DB 2; Length 1236;
Best Local Similarity 34.4%; Pred. No. 7.9;
Matches 11; Conservative 6; Mismatches 9; Indels 6; Gaps 1;

QY 3 AHOHSMETRPDINPAWYAG-----RGIRPV 28
DB 1112 SEQVALETRPMLNPWYAGMLEHGEYGRQI 1143

RESULT 2
Sun/nucleolar protein family protein VC1502 [imported] - Vibrio cholerae (group O1 st
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 15-Sep-2000
C:Accession: A82193
R:Heidelberg, J.F.; Eissen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.
Chardon, D.; Esmolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers
I., R.R.; Mekalanos, J.T.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833
A:Accession: A82193
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-503 <HEP>
A:Cross-references: GB:AE004228; GB:AE003952; NID:95653997; PDB:AA94657.1; GSPDB:GN
A:Experimental source: serogroup O1, strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC1502
A:Map position: 1


```

A:Experimental source: clone F17E5
R:McMurray, A.
Submitted to the EMBL Data Library, November 1995
A:Reference number: Z19378
A:Accession: U21124
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-294 <MT2>
A:Cross-references: PDB:268115; PDB:CAA92169.1; GSPDB:GN00028; CESP:P19H6.1
A:Experimental source: clone F19H6
C:Genetics:
A:Gene: CESP:F19H6.1
A:Map position: X
A:Introns: 55/3; 85/2; 113/3; 150/1; 192/3; 266/2
C:Superfamily: kinase-related transforming protein; protein kinase homology

Query Match      29.7%; Score 51; DB 2; Length 294;
Best Local Similarity 44.8%; Pred. No. 6.3;
Matches 13; Conservative 2; Mismatches 12; Indels 2; Gaps 1;

Oy      3  AHOHSEIRPDINPA--WYAGRGIRPVG 29
      ||| | | | | | | | | | | | | |
Db      139 AHMSKRIRHMDIKPANYFITGNGIVKLG 167

RESULT  8
conserved hypothetical protein MTH1070 - Methanobacterium thermoautotrophicum (strain Delta H)
C:Species: Methanobacterium thermoautotrophicum
C:Date: 29-Jan-1999 #sequence_revision 29-Jan-1999 #text_change 23-Jul-1999
C:Accession: B69009
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Altridge, T.;
; Qiu, D.; Spadafora, R.; Vicalre, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Ulanzi, N.;
; K., S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Neolling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: functi
A:Reference number: A69000; MUID:98037514
A:Accession: B69009
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-430 <MT8>
A:Cross-references: GB:AE000677; GB:AE000666; NID:g2622157; PIDN:AA85559.1; PID:g262214
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH1070
C:Superfamily: conserved hypothetical protein MTH1070

Query Match      29.7%; Score 51; DB 1; Length 430;
Best Local Similarity 69.2%; Pred. No. 9.8;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy      14 DINPANTYAGRGIR 26
      ||| | | | | | | |
Db      191 DINPEWVAGRACR 203

RESULT  9
probable unr protein - rat
S11210
C:Species: Rattus norvegicus (Norway rat)
C:Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 05-Nov-1999
C:Accession: S11210
R:Jeffers, M.; Paciucci, R.; Pellicer, A.
Nucleic Acids Res. 18, 4891-4899, 1990
A:Title: Characterization of unr, a gene closely linked to N-ras.
A:Reference number: S11210; MUID:90570473
A:Accession: S11210
A:Molecule type: mRNA
A:Residues: 1-798 <REP>
A:Cross-references: EMBL:A52311; NID:g57454; PIDN:CAA36549.1; PID:g57455
C:Keywords: DNA binding

```

```

Query Match          29.4%; Score 50.5; DB 2; Length 790;
Best Local Similarity 39.3%; Pred. No. 23;
Matches 10; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY      6 HSMETRPDINPAMYAGRGIRPV 28
      ||: | | | | | | | | | |
Db      583 HSNVGTTEANPTTYSKVRIRPL 605

RESULT 10
H69157
excluclease ABC chain A - Methanobacterium thermoautotrophicum (strain Delta H)
N:Alternate names: uvra protein
M:Contains: excision endonuclease ABC (EC 3.1.1.-) chain A
C:Species: Methanobacterium thermoautotrophicum
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 12-Nov-1999
G:Accession: H69157
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.
J.; Qiu, D.; Spadafora, R.; Vitcalire, R.; Wang, Y.; Wierdowski, J.; Gibson, R.; Jivan,
K.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fu
A:Reference number: A69000; MUID:98037514
A:Accession: H69157
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-962 <MTR>
A:Cross-references: GB:A6000828; GB:A6000666; NID:92621504; PIDN:AMB4949.1; PID:9262
C:Genetics:
A:Experimental source: strain Delta H
A:Gene: MTR443
A:Start codon: TTG
C:Superfamily: excluclease ABC chain A; ATP-binding cassette homology
C:Keywords: ATP; DNA binding; DNA repair; duplication; hydrolase; P-loop
E:38-45/Region: nucleotide-binding motif A (P-loop)
F:632-915/Domain: ATP-binding cassette homology <ABCE>
F:649-656/Region: nucleotide-binding motif A (P-loop)

Query Match          29.7%; Score 51; DB 2; Length 962;
Best Local Similarity 39.5%; Pred. No. 23;
Matches 15; Conservative 0; Mismatches 5; Indels 18; Gaps 2;

QY      11 RPPDINPAMYAG-----RCIRPVGRF 31
      || | | | | | | | | | |
Db      703 RPPRSNPATYGVFTHIRLEFAOTPEARRGRVP-GRF 739

RESULT 11
T47959
hypothetical protein F15G16_60 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
G:Accession: T47959
R:de Haan, M.; Maarse, A.C.; Grivell, L.A.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Q
submitted to the Protein Sequence Database, January 2000
A:Reference number: Z24480
A:Accession: T47959
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-790 <DEH>
A:Cross-references: EMBL:AL132959
A:Experimental source: cultivar Columbia; BAC clone F15G16
C:Genetics:
A:Map position: 3
A:Introns: 39/1; 678/2; 698/3; 773/2
A:Note: F15G16_60

Query Match          29.4%; Score 50.5; DB 2; Length 790;
Best Local Similarity 39.3%; Pred. No. 23;

```

	Matches	11; Conservative	4; Mismatches	12; Indels	1; Gaps
QY	4	HQHSMEIRTPDINPAMYAGRGIRPYGRF	31		
		:			
Db	358	HESYMGFAPPH-NPRTYSGKGLQPHGRW	364		

RESULT 12
T33175
hypothetical protein C24G6.6 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T33175
R:Greco, T.; Bradshaw, H.; Keppler, D.
submitted to the EMBL Data Library, May 1998
A:Description: The sequence of C. elegans cosmid C24G6.
;Reference number: Z21298

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-527 <GB>
A:Cross-references: EMBL:AF067936; PIRN:AA019213.1; GSPBD:GM00023; CESP:C24G6.6
A:Experimental source: strain Bristol N2, clone C24G6
C:Gene(s)
A:Gene: CESP:C24G6.6
A:Map position: 5
A:Introns: 20/3; 77/1; 129/2; 208/3; 470/2

		Query Match	29.1%	Score 50;	DB 2;	Length 527;	
		Best Local Similarity	66.7%	Pred. No. 17;			
		Matches	8;	Conservative	2;	Mismatches	2;
						Indels	0;
						Gaps	0;
QY	13	PDINPMWYAGRG	24				
	1:::						
Db	370	PVYLISMATARG	381				

hypothetical protein - *Synechocystis* sp. (strain PCC 6803)
C/Species: *Synechocystis* sp.
A/Variety: PCC 6803
C/Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
C/Accession: S76955
R./K.: Okumura, T.; Sato, S.; Korani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; Ohta, K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A/Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*
S.
A/Reference number: S74322; MUID: 97061201
A/Accession: S76055

A: accession: preliminary
A: molecule type: DNA
A: residues: 1-128 <KAM>
A: cross-references: EMBL:D90917; GB:AB001339; NID:1655836; PTDN:BAAL8867.1; PTD:d101966
A: note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C:Accession: S76814
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasu, D. K. Res. 3, 103-136, 1996
A>Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* s. s.
A:Reference number: S74322; MUID:97061201
A:Accession: S76814
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-254 <KAN>
A:Cross-references: EMBL:D09016; GB:AB001339; NID:31653715; PTDN:BA018726.1; PID:g1655
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Superfamily: yeast SOL3 protein

Query Match 28.2%; Score 48.5; DB 2; Length 254;
Best Local Similarity 42.3%; Pred. No. 13;
Matches 11; Conservative 5; Mismatches 9; Indels 1; Gaps 1;

	Query Match	28.2%;	Score 48.5;	DB 2;	Length 254;
	Best Local Similarity	42.3%;	Pred. No. 13;		
	Matches 11; Conservative	5;	Mismatches	9;	Indels 1; Gaps 1
Oy	5 QHSM-ETRPDINPAMWAGGIRPVG	29			
	: :	:	: :		
Db	213 QHALGETFAEADPPQPTPARFIOPQG	226			

QY	5	QHSM-BIRTPDINPAMYAGRGIRPVG	29
		:: :::	
DB	213	QHAGGEIFAPEADPQGYPAFTIOPG	218
RESULT	15		

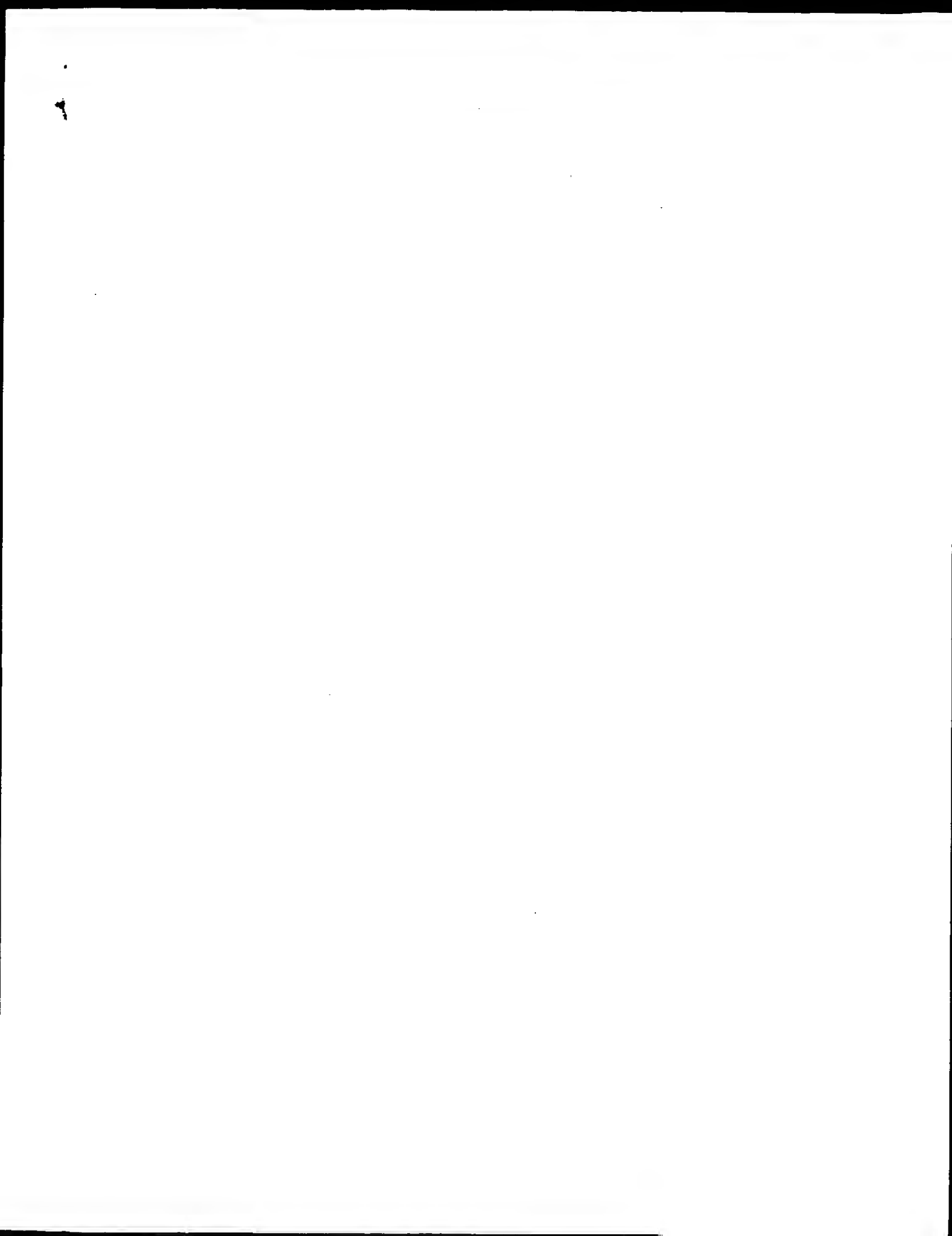
hypothetical protein F8J2.80 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: U475548
R:Wakatsuki, G.; Farmann, B.; Dauner, D.; Sterr, W.; Holland, R.; Weichselgartner, M.
Mayer, K.F.X.
submitted to the Protein Sequence Database, April 2000
.:Reference number: 224458

Search completed: April 17, 2001, 15:45:51
Job time: 599 sec

Tue Apr 17 15:46:14 2001

us-09-446-543a-5.rpt

Page 5



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 17, 2001, 15:38:41, Search time 116.94 Seconds
(without alignments)
9.777 Million cell updates/sec

Title: US-09-446-543a-8

Perfect score: 115

Sequence: 1 TPDINPAMTAGRGIRPVGRF 20

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: A_Geneseq_0401:*

1: /SIDSL/gcgdata/geneseq/geneseq/AA1980.DAT:*
2: /SIDSL/gcgdata/geneseq/geneseq/AA1981.DAT:*
3: /SIDSL/gcgdata/geneseq/geneseq/AA1982.DAT:*
4: /SIDSL/gcgdata/geneseq/geneseq/AA1983.DAT:*
5: /SIDSL/gcgdata/geneseq/geneseq/AA1984.DAT:*
6: /SIDSL/gcgdata/geneseq/geneseq/AA1985.DAT:*
7: /SIDSL/gcgdata/geneseq/geneseq/AA1986.DAT:*
8: /SIDSL/gcgdata/geneseq/geneseq/AA1987.DAT:*
9: /SIDSL/gcgdata/geneseq/geneseq/AA1988.DAT:*
10: /SIDSL/gcgdata/geneseq/geneseq/AA1989.DAT:*
11: /SIDSL/gcgdata/geneseq/geneseq/AA1990.DAT:*
12: /SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT:*
13: /SIDSL/gcgdata/geneseq/geneseq/AA1992.DAT:*
14: /SIDSL/gcgdata/geneseq/geneseq/AA1993.DAT:*
15: /SIDSL/gcgdata/geneseq/geneseq/AA1994.DAT:*
16: /SIDSL/gcgdata/geneseq/geneseq/AA1995.DAT:*
17: /SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT:*
18: /SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT:*
19: /SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT:*
20: /SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT:*
21: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT:*
22: /SIDSL/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	115	100.0	20	W31374	Bovine G protein-c
2	115	100.0	20	W97232	Bovine pituitary-d
3	115	100.0	20	W95191	Bovine pituitary-d
4	115	100.0	20	B10350	Bovine oxytocin se
5	115	100.0	20	Y49301	19P2 ligand peptid
6	115	100.0	21	W31375	Bovine G protein-c
7	115	100.0	21	W95192	Bovine pituitary-d
8	115	100.0	21	B10351	Bovine oxytocin se
9	115	100.0	22	W31376	Bovine G protein-c
10	115	100.0	22	W95193	Bovine pituitary-d
11	115	100.0	22	B10352	Bovine oxytocin se

12	115	100.0	31	W31371	Bovine G protein-c
13	115	100.0	31	W97218	Bovine pituitary-d
14	115	100.0	31	W87613	Bovine 19P2 ligand
15	115	100.0	31	W95188	Bovine pituitary-d
16	115	100.0	31	B10347	Bovine oxytocin se
17	115	100.0	31	Y49290	19P2 ligand peptid
18	115	100.0	31	Y49298	19P2 ligand peptid
19	115	100.0	32	W31372	Bovine G protein-c
20	115	100.0	32	W95189	Bovine pituitary-d
21	115	100.0	32	B10348	Bovine oxytocin se
22	115	100.0	33	W31373	Bovine G protein-c
23	115	100.0	33	W95190	Bovine pituitary-d
24	115	100.0	33	B10349	Bovine oxytocin se
25	115	100.0	33	Y49297	19P2 ligand peptid
26	115	100.0	38	W31382	Bovine genome deriv
27	115	100.0	38	W31368	Bovine G protein-c
28	115	100.0	38	W97224	Bovine genome deriv
29	115	100.0	38	W97217	Bovine pituitary-d
30	115	100.0	38	W95187	Bovine genome deriv
31	115	100.0	98	B10346	Bovine oxytocin se
32	115	100.0	98	B10353	Bovine oxytocin se
33	115	96.5	20	W31387	Rat type G protein
34	115	96.5	20	W97234	Rat type ligand po
35	115	96.5	20	W95175	Murine pituitary-d
36	115	96.5	20	B10358	Rat oxytocin secre
37	115	96.5	20	Y49302	19P2 ligand peptid
38	115	96.5	21	W31388	Rat type G protein
39	115	96.5	21	B10359	Rat oxytocin secre
40	115	96.5	22	W31389	Rat type G protein
41	115	96.5	22	B10360	Rat oxytocin secre
42	115	96.5	31	W31384	Rat type G protein
43	115	96.5	31	W97233	Rat type ligand po
44	115	96.5	31	W87614	Rat 19P2 ligand
45	115	96.5	31	Y87504	Rat prolactin-rele

ALIGNMENTS

RESULT 1	
ID W31374	W31374 standard; Peptide, 20 AA.
XX	W31374;
AC	06-APR-1998 (first entry)
XX	
DE	Bovine G protein-coupled receptor ligand peptide fragment 4.
XX	
KM	G protein-coupled receptor; ligand binding; pharmaceutical;
KW	modulator; pituitary; central nervous system; pancreas; prophylactic;
XX	therapeutic agent.
OS	Bos taurus.
XX	
PN	W09724436-A2.
XX	
PD	10-JUL-1997.
XX	
PF	26-DEC-1996; 96WO-JP03821.
XX	
PR	18-SEP-1996; 96UP-0246573.
PR	28-DEC-1995; 95UP-0343371.
PR	15-MAR-1996; 96UP-0059419.
PR	12-AUG-1996; 96UP-0211805.
XX	
PA	(TAKE) TAKEDA CHEM IND LTD.
XX	
PI	Fuji R, Fukushima S, Habeta Y, Hinuma S, Hosoya M,
PI	Kawamata Y, Kitada C;
DR	WPI: 1997-363672/33.
DR	N-PSDB; V02397.

XX Lligand peptide for G protein-coupled receptor - acts by modulating
 PT function in the central nervous system, pancreas and pituitary gland
 XX
 PS Claim 2; Page 161; 258pp; English.

CC This sequence represents a peptide fragment of a novel bovine pituitary
 CC derived ligand corresponding to amino acid residues 34 to 53 of the
 CC sequence in W31368 and is used in an assay to monitor ligand binding to
 CC the G protein-coupled receptor protein. Pharmaceutical compositions
 CC containing this ligand may be used as a pituitary function modulator. A
 CC central nervous system modulator or a pancreatic function modulator. This
 CC ligand could have specific applications as a prophylactic or therapeutic
 CC agent for dementia, depression, hyperkinetic syndrome, disturbance of
 CC consciousness, anxiety syndrome, schizophrenia, trauma, growth hormone
 CC secretory disease, hyper- and polyphagia, hyperprolactinemia, diabetes,
 CC hyperglycemia, hyperlipidemia, hyperprolactinemia, diabetes,
 CC cancer, pancreatitis, renal disease, Turner's syndrome, neurosis,
 CC rheumatoid arthritis, spinal injury, transient brain ischaemia,
 CC amyotrophic lateral sclerosis, acute myocardial infarction,
 CC spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis,
 CC osteoporosis, asthma, epilepsy, infertility and/or oligosaccharia. Assays
 CC can also be developed to screen compounds which are capable of altering
 CC the binding activity of the ligand thus affecting activation of the G
 CC protein-coupled receptor protein.

CC Sequence 20 AA;

SO

Query Match 100.0%; Score 115; DB 18; Length 20;
 Best Local Similarity 100.0%; Pred. No. 7.7e-12;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDIINPAMYAGRGIRPVGRF 20
 |||||
 DB 1 tpdinpwagrgirpvgrf 20

RESULT 2
 W97232 W97232 standard; peptide; 20 AA.
 ID W97232
 AC W97232;

DT 06-MAY-1999 (first entry)

DE Bovine pituitary-derived ligand polypeptide fragment.

XX Bovine pituitary-derived ligand; modulation; prolactin secretion;
 KW G protein-coupled receptor; GPCR; hypovarianism; gonocyst caecogenesis;
 KW menopause; syndrome; eutypoid; hypometabolism; lactation;
 KW pituitary adenomatosis; brain tumour; amenorrhoea; galactorrhea;
 KW prolactinoma; infertility; impotence; amenorrhoea; galactorrhea;
 KW acromegaly; Chari-Frommel syndrome; Argonz-del Castillo syndrome;
 KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dysospermia;
 KW contraceptive; placental function; chorioncarcinoma; hydatid mole;
 KW interruption mole; abortion; unthrifty fetus; abnormal saccharometabolism;
 KW abnormal lipidmetabolism; oxytocia.

OS Bos sp.

PN W09858962-A1.

PD 30-DEC-1998.

PF 22-JUN-1998; 98WO-JP02765.

PR 23-JUN-1997; 97JP-0165437.

PA (TAKE) TAKEDA CHEM IND LTD.

PI Fujii R, Hinuma S, Kawamata Y, Matsumoto H;

DR WPI; 1999-105614/09.

XX Use of G protein-coupled receptor ligands - for modulating prolactin
 PT secretion or placental function, e.g. for treating menopausal
 PT syndrome, tumours, autoimmune disease or abnormal pregnancy
 XX
 PS Claim 3; Page 136; 241pp; English.

CC The present sequence represents a bovine pituitary-derived ligand
 CC fragment. It is used in the course of the invention. The specification
 CC describes an agent for modulating prolactin secretion which comprises a
 CC ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)
 CC protein. The agents for promoting prolactin secretion can be used for
 CC treating or preventing hypovarianism, gonocyst caecogenesis, menopausal
 CC syndrome, eutypoid or hypometabolism. They can be used for promoting
 CC lactation in a domestic mammal and as an aphrodisiac. The agents for
 CC inhibiting prolactin secretion can be used for treating or preventing
 CC pituitary adenomatosis, brain tumour, amenorrhoea, autoimmune disease,
 CC prolactinoma, infertility, impotence, amenorrhoea, galactorrhea,
 CC acromegaly, Chari-Frommel syndrome, Argonz-del Castillo syndrome,
 CC Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dysospermia.
 CC The inhibitory agents can also be used as contraceptives. The agents for
 CC modulating placental function can be used for treating or preventing
 CC chorioncarcinoma, hydatid mole, interruption mole, abortion, unthrifty fetus,
 CC abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.

SO Sequence 20 AA;

Query Match 100.0%; Score 115; DB 20; Length 20;
 Best Local Similarity 100.0%; Pred. No. 7.7e-12;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDIINPAMYAGRGIRPVGRF 20
 |||||
 DB 1 tpdinpwagrgirpvgrf 20

RESULT 3
 W95191 W95191 standard; peptide; 20 AA.
 ID W95191
 AC W95191;

DT 10-MAR-1999 (first entry)

DE Bovine pituitary-derived ligand polypeptide fragment.

XX pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;
 KW GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas;
 KW tissue; screen; therapeutic; binding; senile dementia; ligand; murine;
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;
 KW Creutzfeldt-Jakob disease; poisoning; schizophrenia; growth hormone;
 KW secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;
 KW gene therapy; transgenic animal; bovine.

OS Bos sp.

PN W09849295-A1.

PD 05-NOV-1998.

PF 27-APR-1998; 98WO-JP01923.

PR 28-APR-1997; 97JP-0109974.

PA (TAKE) TAKEDA CHEM IND LTD.

PI Fukusumi S, Hinuma S;

WPI; 1999-009423/01.

PT New polypeptide ligand for orphan G protein coupled receptors - used

PT for treating disorders of central nervous system, pituitary and
PT pancreas, and for drug screening

Example 19, Page 151; 206pp; English.

CC The invention relates to a murine pituitary-derived ligand polypeptide
CC which is a ligand for the G-protein coupled orphan receptor designated
CC GPR10 (human) or UMR-1 (rat). Cells transformed with a vector containing
CC the ligand polypeptide encoding DNA are used to produce a recombinant
CC ligand polypeptide. The ligand polypeptide, and its fragments, modulate
CC function of the pituitary, central nervous system, pancreas and other
CC tissues and can be used to screen for agents that modulate binding of the
CC polypeptide to the receptor, to quantify the amount of receptor in a
CC sample and to raise antibodies. They may also be used therapeutically,
CC e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's
CC diseases; Creutzfeldt-Jakob disease; poisoning by heavy metals or drugs;
CC diabetes; schizophrenia; disorders of growth hormone secretion; cancer;
CC rheumatoid arthritis, epilepsy and many others, also to improve post-
CC operative nutritional status and as vasopressor. Transgenic animals
CC carrying the ligand polypeptide encoding DNA or its mRNA are used to
CC study the function of the polypeptide-expressing genes, as models of
CC disease, for drug screening and as source of cell lines. The ligand
CC polypeptide DNA is used as a source of probes and primers; to identify
CC related sequences; in receptor-binding assays; for production of Ab and
CC antisera; in drug development; for gene therapy and to develop transgenic
CC animals. The present sequence represents a bovine genome-derived ligand
CC polypeptide fragment which is similar to the murine ligand-polypeptide.

SO Sequence 20 AA;

Query Match 100.0%; Score 115; DB 20; Length 20;
Best Local Similarity 100.0%; Pred. No. 7,7e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAMYAGRGIRPVGRF 20
| | | | | | | | | | | | | | | | | | | | | |
Db 1 tpdinpwagrgirpvgrrf 20

RESULT 4

ID B10350 standard; peptide; 20 AA.

AC B10350;

DT 24-NOV-2000 (first entry)

DE Bovine oxytocin secretion promoting peptide SEQ ID NO: 6.

KM Bovine; oxytocin secretion promoter; G protein-coupled receptor protein;
KM treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;
KM caesarean section; artificial fertilization; galactostasis; goat; pig;
KM veterinary medicine; milk production.

OS Bos taurus.

PN WO200038704-A1.

PD 06-JUL-2000.

PF 22-DEC-1999; 99MO-JP07199.

PR 25-DEC-1998; 98JP-0369585.

PA (TAKE) TAKEDA CHEM IND LTD.

PI Matsumoto H, Kitada C, Hinuma S;

DR WPI; 2000-452298/39.

PT Physiologically-active polypeptide recognized as ligand by G
PT protein-coupled receptor protein, for promoting secretion of oxytocin,

PT as drugs for diseases relating to oxytocin secretion and in veterinary
PT medicine

PS Claim 5; Page 51; 72pp; Japanese.

CC This invention describes a novel oxytocin secretion-regulating agent
CC which contains a ligand peptide or its salt for the G protein-coupled
CC receptor protein. It is useful in the form of drugs for ameliorating,
CC preventing and treating diseases relating to oxytocin secretion e.g.,
CC weak pains and atonic bleeding, before and after expulsion of placenta,
CC uterine recovery failure, caesarean section, stoppage of artificial
CC fertilization or galactostasis and is also applicable in veterinary
CC medicine for promoting milk production in cow, goat and pig. This
CC sequence represents a bovine peptide which acts as an oxytocin secretion
CC promoter.

SO Sequence 20 AA;

Query Match 100.0%; Score 115; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 7,7e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAMYAGRGIRPVGRF 20
| | | | | | | | | | | | | | | | | | | | | |
Db 1 tpdinpwagrgirpvgrrf 20

RESULT 5

ID Y49301 standard; peptide; 20 AA.

AC Y49301;

DT 22-FEB-2000 (first entry)

DE 19P2 ligand peptide fragment.

KM Monoclonal antibody; 19P2 ligand; diagnosis; prolactin secretion;
KM pituitary; regulatory mechanism; central nervous system; pancreatic.

OS Bos sp.

PN Key Location/Qualifiers

PT Modified-site 20 /note= "C-terminal amide"

PD WO960112-A1.

PF 25-NOV-1999.

PR 20-MAY-1999; 99MO-JP02650.

PA 21-MAY-1998; 98JP-0140293.

PI (TAKE) TAKEDA CHEM IND LTD.

PN Matsumoto H, Kitada C, Hinuma S;

DR WPI; 2000-039381/03.

PT New monoclonal antibodies, useful in diagnosis, as drugs and in
PT studying diseases related to ligand abnormality
PS Disclosure; Page 27; 73pp; Japanese.

CC The invention provides a monoclonal antibody which has a specific
CC reaction with the part peptide of the C-terminal of 19P2 ligand or its
CC derivative. The antibodies can be used in diagnosis or to treat or
CC prevent diseases associated with abnormality in the pituitary function
CC regulatory mechanism (e.g. promotion of prolactin secretion), central
CC nervous regulatory mechanism, and pancreatic function regulatory
CC mechanism. The antibody-based immunoassay can also be applied in

CC clarifying the physiological functions of the ligand and its derivative.
 CC Sequences Y49290-302 represent peptide fragments of the 19p2 ligand.
 XX
 SQ Sequence 20 AA;

Query Match 100.0%; Score 115; DB 21; Length 20;
 Best Local Similarity 100.0%; Pred. No. 7.7e-12;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TPDIINPAMYAGRGIRPYGRF 20
 |||
 DB 1 tpdinpawyagrgrirpygrf 20

RESULT 6
 W31375
 ID W31375 standard; Peptide; 21 AA.
 AC W31375;
 XX
 XX 06-APR-1998 (first entry)

DE Bovine G protein-coupled receptor ligand peptide fragment 5.
 XX
 XX

KW G protein-coupled receptor; ligand binding; pharmaceutical;
 KW modulator; pituitary; central nervous system; pancreas; prophylactic;
 KW therapeutic agent.

OS Bos taurus.

XX W05724436-A2.

XX 10-JUL-1997.

XX 26-DEC-1996; 96MO-JP03821.

XX 18-SEP-1996; 96JP-0246573.

XX 28-DEC-1995; 95JP-0343371.

XX 15-MAR-1996; 96JP-0059419.

XX 12-AUG-1996; 96JP-0211805.

XX (TAKE) TAKEDA CHEM IND LTD.

XX Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;
 PI Kawamata Y, Kitada C;
 XX WPI: 1997-363672/33.
 DR N-PSDB; V02398.

XX Ligand peptide for G protein-coupled receptor - acts by modulating
 PT function in the central nervous system, pancreas and pituitary gland
 XX
 XX Claim 2; Page 162; 258pp; English.

XX This sequence represents a peptide fragment of a novel bovine pituitary
 CC derived ligand corresponding to amino acid residues 34 to 54 of the
 CC sequence in W31375 and is used in an assay to monitor ligand binding to
 CC the G protein-coupled receptor protein. Pharmaceutical compositions
 CC containing this ligand may be used as a pituitary function modulator, a
 CC central nervous system modulator or a pancreatic function modulator. This
 CC ligand could have specific applications as a prophylactic or therapeutic
 CC agent for dementia, depression, hyperkinetic syndrome, disturbance of
 CC consciousness, anxiety syndrome, schizophrenia, trauma, growth hormone
 CC secretory disease, hyper- and polypagia, hypercholesterolaemia,
 CC hyperglycaemia, hyperlipidaemia, hyperprolactinaemia, diabetes,
 CC cancer, pancreatitis, renal disease, Turner's syndrome, neurosis,
 CC rheumatoid arthritis, spinal injury, transient brain ischaemia,
 CC amyloidotic lateral sclerosis, acute myocardial infarction,
 CC spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis,
 CC osteoporosis, asthma, epilepsy, infertility and/or oligosactia. Assays
 CC can also be developed to screen compounds which are capable of altering
 CC the binding activity of the ligand thus affecting activation of the G

CC protein-coupled receptor protein.

XX
 SQ Sequence 21 AA;

Query Match 100.0%; Score 115; DB 18; Length 21;
 Best Local Similarity 100.0%; Pred. No. 8.1e-12;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TPDIINPAMYAGRGIRPYGRF 20
 |||
 DB 1 tpdinpawyagrgrirpygrf 20

RESULT 7
 W95192
 ID W95192 standard; peptide; 21 AA.
 AC W95192;
 XX
 XX 10-MAR-1999 (first entry)

DE Bovine pituitary-derived ligand polypeptide fragment.
 XX
 XX

KW Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;
 KW GPR10; modulator; pituitary; central nervous system; pancreas;
 KW tissue; screen; therapeutic; binding; senile dementia; ligand; murine;
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;
 KW Creutzfeld-Jakob disease; poisoning; schizophrenia; growth hormone;
 KW secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;
 KW gene therapy; transgenic animal; bovine.

OS Bos sp.

XX W09849295-A1.

XX 05-NOV-1998.

XX 27-APR-1998; 98MO-JP01923.

XX 28-APR-1997; 97JP-0109974.

XX (TAKE) TAKEDA CHEM IND LTD.

XX Fukusumi S, Hinuma S;
 PI
 XX WPI: 1999-009423/01.

XX New polypeptide ligand for orphan G protein coupled receptors - used
 PT for treating disorders of central nervous system, pituitary and
 PT pancreas, and for drug screening
 XX
 XX Example 19; Page 151; 206pp; English.

XX The invention relates to a murine pituitary-derived ligand polypeptide
 CC which is a ligand for the G-protein coupled orphan receptor designated
 CC GPR10 (human) or GHR-1 (rat). Cells transformed with a vector containing
 CC the ligand polypeptide encoding DNA are used to produce a recombinant
 CC ligand polypeptide. The ligand polypeptide, and its fragments, modulate
 CC function of the pituitary, central nervous system, pancreas and other
 CC tissues and can be used to screen for agents that modulate binding of the
 CC polypeptide to the receptor; to quantify the amount of receptor in a
 CC sample and to raise antibodies. They may also be used therapeutically,
 CC e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's
 CC diseases; Creutzfeld-Jakob disease; poisoning by heavy metals or drugs;
 CC diabetes; schizophrenia; disorders of growth hormone secretion; cancer;
 CC rheumatoid arthritis, epilepsy and many others, also to improve post-
 CC operative nutritional status and as vasopressor. Transgenic animals
 CC carrying the ligand polypeptide encoding DNA or its mutant are used to
 CC study the function of the polypeptide-expressing genes, as models of
 CC disease, for drug screening and as source of cell lines. The ligand
 CC polypeptide DNA is used as a source of probes and primers; to identify
 CC related sequences; in receptor-binding assays; for production of Ab and

CC antisera; in drug development; for gene therapy and to develop transgenic
 CC animals. The present sequence represents a bovine genome-derived ligand
 CC polypeptide fragment which is similar to the murine ligand-polypeptide.

XX Sequence 21 AA;

Query Match 100.0%; Score 115; DB 20; Length 21;
 Best Local Similarity 100.0%; Pred. No. 8, 1e-12;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRGIRPVGRF 20
 |||||
 Db 1 tpdinpaawyagrgirpvgrf 20

RESULT 8

ID B10351 standard; peptide: 21 AA.

AC B10351;

DT 24-NOV-2000 (first entry)

DE Bovine oxytocin secretion promoting peptide SEQ ID NO: 7.

KW Bovine; oxytocin secretion promoter; G protein-coupled receptor protein;
 KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;
 KW caesarean section; artificial fertilization; galactostasis; goat; pig;
 KW veterinary medicine; milk production.

OS Bos taurus.

PN W0200038704-A1.

PD 06-JUL-2000.

PE 22-DEC-1999; 99WO-JP07199.

PR 25-DEC-1998; 98JP-0369585.

PA (TAKE) TAKEDA CHEM IND LTD.

PI Matsumoto H, Kitada C, Hinuma S;

DR WPI; 2000-452298/39.

PT Physiologically-active polypeptide recognized as ligand by G

PT protein-coupled receptor protein, for promoting secretion of oxytocin,

PT as drugs for diseases relating to oxytocin secretion and in veterinary

PT medicine -

PS Disclosure; Page 52; 72pp; Japanese.

CC This invention describes a novel oxytocin secretion-regulating agent
 CC which contains a ligand peptide or its salt for the G protein-coupled
 CC receptor protein. It is useful in the form of drugs for ameliorating,
 CC preventing and treating diseases relating to oxytocin secretion e.g.
 CC weak pains and atonic bleeding, before and after expulsion of placenta,
 CC uterine recovery failure, caesarean section, stoppage of artificial
 CC fertilization or galactostasis and is also applicable in veterinary
 CC medicine for promoting milk production in cow, goat and pig. This
 CC sequence represents a bovine peptide which acts as an oxytocin secretion
 CC promoter.

XX Sequence 21 AA;

Query Match 100.0%; Score 115; DB 21; Length 21;
 Best Local Similarity 100.0%; Pred. No. 8, 1e-12;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRGIRPVGRF 20

Db 1 tpdinpaawyagrgirpvgrf 20

RESULT 9

ID W31376 standard; Peptide: 22 AA.

AC W31376;

DT 06-APR-1998 (first entry)

DE Bovine G protein-coupled receptor ligand peptide fragment 6.

KW G protein-coupled receptor; ligand binding; pharmaceutical;
 KW modulator; pituitary; central nervous system; pancreas; prophylactic;
 KW therapeutic agent.

OS Bos taurus.

PN W09724436-A2.

PD 10-JUL-1997.

PE 26-DEC-1996; 96WO-JP03821.

PR 18-SEP-1996; 96JP-0246573.

PR 28-DEC-1995; 95JP-0143371.

PR 15-MAR-1996; 96JP-0059419.

PR 12-AUG-1996; 96JP-0211805.

PA (TAKE) TAKEDA CHEM IND LTD.

PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;

PI Kawamata Y, Kitada C;

DR WPI: 1997-363672/33.

DR N-PSDB; V023399.

PT Claim 2; Page 162; 258pp; English.

CC This sequence represents a peptide fragment of a novel bovine pituitary
 CC derived ligand corresponding to amino acid residues 34 to 55 of the
 CC sequence in W31368 and is used in an assay to monitor ligand binding to
 CC the G protein-coupled receptor protein. Pharmaceutical compositions
 CC containing this ligand may be used as a pituitary function modulator, a
 CC central nervous system modulator or a pancreatic function modulator. This
 CC ligand could have specific applications as a prophylactic or therapeutic
 CC agent for dementia, depression, hyperkinetic syndrome, disturbance of
 CC consciousness, anxiety syndrome, schizophrenia, trauma, growth hormone
 CC secretory disease, hyper- and polyphagia, hypercholesterolemia,
 CC hyperglycemia, hyperlipidemia, hyperproliferation, diabetes,
 CC cancer, pancreatitis, renal disease, Turner's syndrome, neurosis,
 CC rheumatoid arthritis, spinal injury, transient brain ischemia,
 CC amyotrophic lateral sclerosis, acute myocardial infarction,
 CC spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis,
 CC osteoporosis, asthma, epilepsy, infertility and/or oligogalactia. Assays
 CC can also be developed to screen compounds which are capable of altering
 CC the binding activity of the ligand thus affecting activation of the G
 CC protein-coupled receptor protein.

XX Sequence 22 AA;

Query Match 100.0%; Score 115; DB 18; Length 22;
 Best Local Similarity 100.0%; Pred. No. 8, 5e-12;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYAGRGIRPVGRF 20

DB 1 tpdinpawagrglrpvgri 20

RESULT 10

W95193

ID W95193 standard; peptide: 22 AA.

XX

AC W95193;

XX

DT 10-MAR-1999 (first entry)

XX

DE Bovine pituitary-derived ligand polypeptide fragment.

XX

KM Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor; GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas; tissue; screen; therapeutic; binding; senile dementia; ligand; murine; Alzheimer's disease; Parkinson's disease; Huntington's disease; drug; Creutzfeld-Jakob disease; poisoning; schizophrenia; growth hormone; secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor; gene therapy; transgenic animal; bovine.

KM

XX

OS Bos sp.

XX

PN W09849295-A1.

XX

PD 05-NOV-1998.

XX

PF 27-APR-1998; 98WO-JP01923.

XX

PR 28-APR-1997; 97JP-0109974.

XX

PA (TAKE) TAKEDA CHEM IND LTD.

XX

PI Fukusumi S, Hinuma S;

XX

DR MPI; 1999-009423/01.

XX

PT New polypeptide ligand for orphan G protein coupled receptors - used for treating disorders of central nervous system, pituitary and pancreas, and for drug screening

PT

XX

PS Example 19; Page 151; 206pp; English.

XX

XX The invention relates to a murine pituitary-derived ligand polypeptide which is a ligand for the G-protein coupled orphan receptor designated GPR10 (human) or UHR-1 (rat). Cells transformed with a vector containing the ligand polypeptide encoding DNA are used to produce a recombinant ligand polypeptide. The ligand polypeptide, and its fragments, modulate function of the pituitary, central nervous system, pancreas and other tissues and can be used to screen for agents that modulate binding of the polypeptide to the receptor; to quantify the amount of receptor in a sample and to raise antibodies. They may also be used therapeutically, e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's diseases; Creutzfeld-Jakob disease; poisoning by heavy metals or drugs; diabetes; schizophrenia; disorders of growth hormone secretion; cancer; rheumatoid arthritis; epilepsy and many others, also to improve post-operative nutritional status and as vasopressor. Transgenic animals carrying the ligand polypeptide encoding DNA or its mutant are used to study the function of the polypeptide-expressing genes, as models of disease, for drug screening and as source of cell lines. The ligand polypeptide DNA is used as a source of probes and primers; to identify related sequences; in receptor-binding assays; for production of Ab and antisera; in drug development; for gene therapy and to develop transgenic animals. The present sequence represents a bovine genome-derived ligand polypeptide fragment which is similar to the murine ligand-polypeptide.

XX

XX

Sequence 22 AA;

Query Match 100.0%; Score 115; DB 20; Length 22;

Best Local Similarity 100.0%; Pred. No. 8.5e-12;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAMWAGRGIRPVGRF 20

DB 1 tpdinpawagrglrpvgri 20

RESULT 11

ID B10352

XX B10352 standard; peptide: 22 AA.

XX

AC B10352;

XX

DT 24-NOV-2000 (first entry)

XX

DE Bovine oxytocin secretion promoting peptide SEQ ID NO: 8.

XX

KM Bovine; oxytocin secretion promoter; G protein-coupled receptor protein; treatment; disease; pain; atonic bleeding; uterine recovery failure; cow; caesarean section; artificial fertilization; galactostasis; goat; pig; veterinary medicine; milk production.

KM

XX

OS Bos taurus.

XX

PN W0200038704-A1.

XX

PD 06-JUL-2000.

XX

PF 22-DEC-1999; 99WO-JP07199.

XX

PR 25-DEC-1998; 98JP-0369585.

XX

PA (TAKE) TAKEDA CHEM IND LTD.

XX

PI Matsumoto H, Kitada C, Hinuma S;

XX

DR MPI: 2000-452298/39.

XX

PT Physiologically-active polypeptide recognized as ligand by G protein-coupled receptor protein, for promoting secretion of oxytocin, as drugs for diseases relating to oxytocin secretion and in veterinary medicine -

PT

XX

PS Disclosure; Page 52; 72pp; Japanese.

XX

XX This invention describes a novel oxytocin secretion-regulating agent which contains a ligand peptide or its salt for the G protein-coupled receptor protein. It is useful in the form of drugs for ameliorating, preventing and treating diseases relating to oxytocin secretion e.g. weak pains and atonic bleeding, before and after expulsion of placenta, uterine recovery failure, caesarean section, stoppage of artificial fertilization or galactostasis and is also applicable in veterinary medicine for promoting milk production in cow, goat and pig. This sequence represents a bovine peptide which acts as an oxytocin secretion promoter.

XX

XX

Sequence 22 AA;

QY 1 TPDINPAMWAGRGIRPVGRF 20

DB 1 tpdinpawagrglrpvgri 20

Query Match 100.0%; Score 115; DB 21; Length 22;

Best Local Similarity 100.0%; Pred. No. 8.5e-12;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAMWAGRGIRPVGRF 20

DB 1 tpdinpawagrglrpvgri 20

RESULT 12

ID W31371

XX W31371 standard; Peptide; 31 AA.

XX

AC W31371;

XX

DT 06-APR-1998 (first entry)

Query Match 100.0%; Score 115; DB 20; Length 22;

Best Local Similarity 100.0%; Pred. No. 8.5e-12;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DE Bovine 19P2 ligand.
 XX 19P2 ligand; G protein coupled receptor; pituitary;
 KW prolactin releasing peptide; cattle; dementia; breast cancer;
 XX therapy.
 XX Bos taurus.
 XX EP887417-A2.
 XX 30-DEC-1998.
 XX 25-JUN-1998; 98EP-0111725.
 XX 27-JUN-1997; 97JP-0172118.
 XX (TAKE) TAKEDA CHEM IND LTD.
 XX Moriya T, Nishimura O, Suenaga M, Tanaka Y;
 XX WPI; 1999-047884/05.
 XX Producing a 19P2 pituitary G protein receptor ligand - by cleavage
 PT of a fusion protein, useful for preventing and treating dementia,
 PT breast cancer, renal failure and autoimmune disease
 XX
 XX Claim 5; Page 34; 56pp; English.
 XX
 XX This is the amino acid sequence of the bovine pituitary G
 CC protein-coupled receptor ligand 19P2. A method suitable for
 CC commercial high-level production of 19P2 comprises expressing
 CC the ligand in host cells as a recombinant fusion protein e.g. with
 CC human basic fibroblast growth factor (see V83792-93) that has
 CC been modified to include an N-terminal cysteine residue. The
 CC ligand is released from the fusion by cyanidation followed by
 CC ammonolysis. 19P2 has prolactin secretion-stimulating and (at
 CC high doses) prolactin secretion-inhibiting properties. It can be
 CC used in the treatment and prevention of various diseases including:
 CC senile dementia, cerebrovascular dementia, and dementia associated
 CC with: neurological disorders (e.g. Alzheimer's disease, Parkinson's
 CC disease, Pick's disease, Huntington's disease), infectious diseases
 CC (e.g. Creutzfeldt-Jakob's), endocrine or metabolic disease or
 CC toxicosis (e.g. hypothyroidism, vitamin B12 deficiency, alcoholism,
 CC intoxication by drugs, metal and organic compounds), tumorigenic
 CC diseases (e.g. brain tumour), traumatic diseases (e.g. chronic
 CC subarachnoid haemorrhage), and other types of dementia, depression,
 CC hyperactive child syndrome (microencephalopathy) and disturbance of
 CC consciousness. It is also useful for prevention and treatment of
 CC diseases associated with prolactin hypo and hypersecretion
 CC respectively, including: hyperprolactinaemia, pituitary adenoma,
 CC breast cancer, infertility, impotence and autoimmune disease
 CC (hypersecretion disorders), and seminal vesicle hypoplasia,
 CC osteoporosis, menopausal syndrome and renal failure (hyposecretion
 CC disorders). The 19P2 polypeptide/amide is also useful as a test
 CC reagent for study of the prolactin secretory function or as a
 CC lactagogue in mammalian farm animals.
 XX
 SO Sequence 31 AA;

Query Match 100.0%; Score 115; DB 20; Length 31;
 Best Local Similarity 100.0%; Pred. No. 1.2e-11;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDPNPAMYAGRGIRPVGRF 20
 ||||||||||||||||
 DB 12 tpdnpawyagrgirpgrf 31

RESULT 15
 W95188
 ID *W95188-standard; peptide; 31 AA.
 XX

AC W95188;
 XX 10-MAR-1999 (first entry)
 XX Bovine pituitary-derived ligand polypeptide fragment.
 DE
 XX Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;
 KW GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas;
 KW tissue; screen; therapeutic; binding; senile dementia; ligand; murine;
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;
 KW Creutzfeldt-Jakob disease; poisoning; schizophrenia; growth hormone;
 KW secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;
 KW gene therapy; transgenic animal; bovine.
 XX
 XX Bos sp.
 XX OS
 XX NO9849295-A1.
 XX 05-NOV-1998.
 XX 27-APR-1998; 98MO-JF01923.
 XX 28-APR-1997; 97JP-0109974.
 XX (TAKE) TAKEDA CHEM IND LTD.
 XX Fukusumi S, Hinuma S;
 XX WPI; 1999-009423/01.
 XX New polypeptide ligand for orphan G protein coupled receptors - used
 PT for treating disorders of central nervous system, pituitary and
 PT pancreas, and for drug screening
 XX
 XX Example 19; Page 150; 206pp; English.
 XX
 XX The invention relates to a murine pituitary-derived ligand polypeptide
 CC which is a ligand for the G-protein coupled orphan receptor designated
 CC GPR10 (human) or UHR-1 (rat). Cells transformed with a vector containing
 CC the ligand polypeptide encoding DNA are used to produce a recombinant
 CC ligand polypeptide. The ligand polypeptide, and its fragments, modulate
 CC function of the pituitary, central nervous system, pancreas and other
 CC tissues and can be used to screen for agents that modulate binding of the
 CC polypeptide to the receptor; to quantify the amount of receptor in a
 CC sample and to raise antibodies. They may also be used therapeutically,
 CC e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's
 CC diseases; Creutzfeldt-Jakob disease; poisoning by heavy metals or drugs;
 CC diabetes; schizophrenia; disorders of growth hormone secretion; cancer;
 CC rheumatoid arthritis; epilepsy and many others, also to improve post-
 CC operative nutritional status and as vasopressor. Transgenic animals
 CC carrying the ligand polypeptide encoding DNA or its mutant are used to
 CC study the function of the polypeptide-expressing genes, as models of
 CC disease, for drug screening and as source of cell lines. The ligand
 CC polypeptide DNA is used as a source of probes and primers; to identify
 CC related sequences; in receptor-binding assays; for production of Ab and
 CC antisera; in drug development; for gene therapy and to develop transgenic
 CC animals. The present sequence represents a bovine genome-derived ligand
 CC polypeptide fragment which is similar to the murine ligand-polypeptide.
 XX
 SO Sequence 31 AA;

Query Match 100.0%; Score 115; DB 20; Length 31;
 Best Local Similarity 100.0%; Pred. No. 1.2e-11;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDPNPAMYAGRGIRPVGRF 20
 ||||||||||||||||
 DB 12 tpdnpawyagrgirpgrf 31

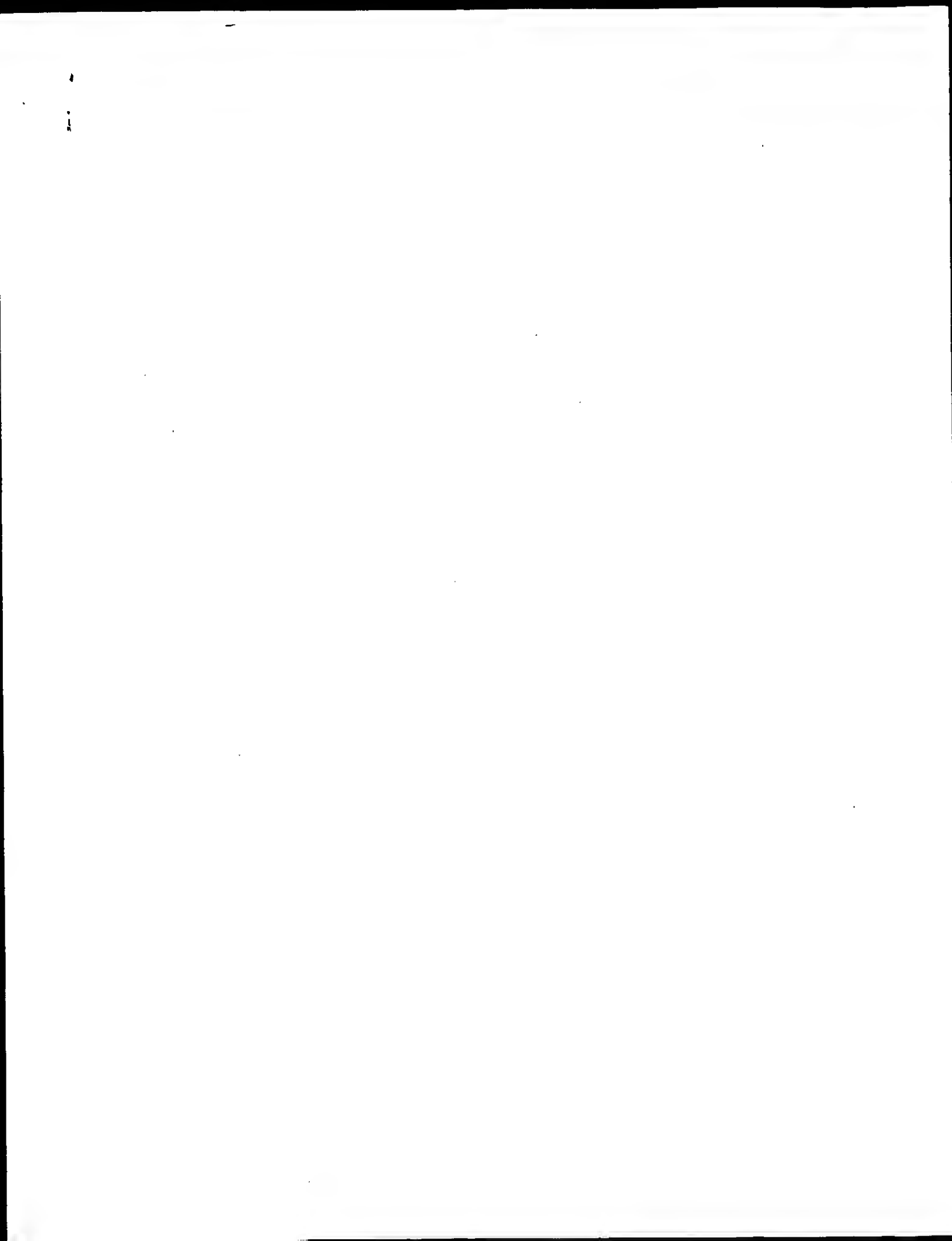
Search completed: April 17, 2001, 15:38:42

Tue Apr 17 15:46:27 2001

us-09-446-543a-8.rag

Page 9

Job time: 349 sec



Tue Apr 17 15:46:27 2001

US-09-446-543a-8.ra1

Page 1

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 17, 2001, 15:39:47 ; Search time 61.54 Seconds

(without alignments)
6.243 Million cell updates/sec

File: US-09-446-543a-8

Perfect score: 115

Sequence: 1 TPDIINPMWYAGRIPIYGRF 20

Scoring table: BIOSUM62

Gapop 10.0, Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents, AA:
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PC105.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfill1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	115	100.0	20	3	US-09-105-678A-34
2	115	100.0	21	3	US-09-105-678A-35
3	115	100.0	22	3	US-09-105-678A-36
4	115	100.0	31	3	US-09-105-678A-7
5	115	100.0	31	3	US-09-105-678A-31
6	115	100.0	32	3	US-09-105-678A-32
7	115	100.0	32	3	US-09-105-678A-33
8	115	100.0	32	3	US-09-105-678A-34
9	115	100.0	32	3	US-09-105-678A-35
10	115	100.0	32	3	US-09-105-678A-36
11	115	100.0	32	3	US-09-105-678A-37
12	115	100.0	32	3	US-09-105-678A-38
13	115	100.0	32	3	US-09-105-678A-39
14	115	100.0	32	3	US-09-105-678A-40
15	115	100.0	32	3	US-09-105-678A-41
16	115	100.0	32	3	US-09-105-678A-42
17	115	100.0	32	3	US-09-105-678A-43
18	115	100.0	32	3	US-09-105-678A-44
19	115	100.0	32	3	US-09-105-678A-45
20	115	100.0	32	3	US-09-105-678A-46
21	115	100.0	32	3	US-09-105-678A-47
22	115	100.0	32	3	US-09-105-678A-48
23	115	100.0	32	3	US-09-105-678A-49
24	115	100.0	32	3	US-09-105-678A-50
25	115	100.0	32	3	US-09-105-678A-51
26	115	100.0	32	3	US-09-105-678A-52
27	115	100.0	32	3	US-09-105-678A-53

28	46	40.0	555	2	US-08-982-232-7	Sequence 7, Appl
29	46	40.0	555	2	US-08-982-232-14	Sequence 14, Appl
30	44	38.3	774	3	US-08-902-632-2	Sequence 2, Appl
31	44	38.3	774	3	US-09-073-354-1	Sequence 1, Appl
32	44	38.3	774	3	US-08-656-005A-1	Sequence 1, Appl
33	44	38.3	774	4	US-09-073-259-1	Sequence 1, Appl
34	44	38.3	774	4	US-09-363-095-1	Sequence 1, Appl
35	44	35.7	70	4	US-08-513-974B-34	Sequence 34, Appl
36	41	35.7	70	4	US-08-513-974B-317	Sequence 317, App
37	41	35.7	141	4	US-08-513-974B-320	Sequence 320, App
38	41	35.7	209	4	US-08-513-974B-321	Sequence 321, App
39	41	35.7	313	2	US-08-446-806-1	Sequence 1, Appl
40	41	35.7	330	2	US-08-815-176-1	Sequence 1, Appl
41	40	34.8	396	1	US-08-229-284A-2	Sequence 2, Appl
42	40	34.8	426	6	5268463-8	Patent No. 5268463
43	40	34.8	428	6	5432081-9	Patent No. 5432081
44	40	34.8	456	6	5432081-7	Patent No. 5432081
45	40	34.8	457	6	5268463-7	Patent No. 5268463

ALIGNMENTS

RESULT 1
US-09-105-678A-34
Sequence 34, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19p2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESSES:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-34

Query Match 100.0%; Score 115; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 5.2e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAMYAGRGIRPVGRF 20
Db 1 TPDINPAMYAGRGIRPVGRF 20

RESULT 2

US-09-105-678A-35
Sequence 35, Application US/09105678A
Patent No. 6103882

GENERAL INFORMATION:

APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 35:

SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-35

Query Match 100.0%; Score 115; DB 3; Length 21;
Best Local Similarity 100.0%; Pred. No. 5.5e-12;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAMYAGRGIRPVGRF 20
Db 1 TPDINPAMYAGRGIRPVGRF 20

RESULT 3

US-09-105-678A-36
Sequence 36, Application US/09105678A
Patent No. 6103882

GENERAL INFORMATION:

APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 36:

SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-36

Query Match 100.0%; Score 115; DB 3; Length 22;
Best Local Similarity 100.0%; Pred. No. 5.8e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DY 1 TPDINPAMYAGRGIRPVGRF 20
Db 1 TPDINPAMYAGRGIRPVGRF 20

RESULT 4

US-09-105-678A-7
Sequence 7, Application US/09105678A
Patent No. 6103882

GENERAL INFORMATION:

APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.

REGISTRATION NUMBER: 27, 026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-7

Query Match 100.0%; Score 115; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 8.5e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TPDINPAMYAGRGIRPYGRF 20
DB 12 TPDINPAMYAGRGIRPYGRF 31

RESULT 5
US-09-105-678A-31
Sequence 31, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27, 026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-31

Query Match 100.0%; Score 115; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 8.5e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAMYAGRGIRPYGRF 20
DB 12 TPDINPAMYAGRGIRPYGRF 31

RESULT 6
US-09-105-678A-32
Sequence 32, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27, 026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 32 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-32

Query Match 100.0%; Score 115; DB 3; Length 32;
Best Local Similarity 100.0%; Pred. No. 8.8e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAMYAGRGIRPYGRF 20
DB 12 TPDINPAMYAGRGIRPYGRF 31

RESULT 7
US-09-105-678A-33
Sequence 33, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:

ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-33

Query Match 100.0%; Score 115; DB 3; Length 33;
Best Local Similarity 100.0%; Pred. No. 9.1e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDIHPAWYAGRGIRPVGRF 20
|||||
DB 12 TPDIHPAWYAGRGIRPVGRF 31

RESULT 8
US-09-105-678A-40
Sequence 40, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
ATTORNEY/AGENT INFORMATION:

NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-41

Query Match 96.5%; Score 111; DB 3; Length 20;
Best Local Similarity 95.0%; Pred. No. 2.2e-11;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPDIHPAWYAGRGIRPVGRF 20
|||||
DB 1 TPDIHPAWYAGRGIRPVGRF 20

RESULT 9
US-09-105-678A-41
Sequence 41, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-41

Query Match 96.5%; Score 111; DB 3; Length 21;
Best Local Similarity 95.0%; Pred. No. 2.4e-11;

Matches 19: Conservative 0: Mismatches 1: Indels 0: Gaps 0:

OY 1 TPDINPAMYAGRGIRPVGRF 20
|||||
DB 1 TPDINPAMYTGIRGVGRF 20

RESULT 10

US-09-105-678A-42
Sequence 42, Application US/09105678A
Patent No. 6103882

GENERAL INFORMATION:

APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-42

Query Match 96.5%; Score 111; DB 3; Length 22;
Best Local Similarity 95.0%; Pred. No. 2.5e-11;
Matches 19: Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TPDINPAMYAGRGIRPVGRF 20
|||||
DB 1 TPDINPAMYTGIRGVGRF 20

RESULT 11

US-09-105-678A-8
Sequence 8, Application US/09105678A
Patent No. 6103882

GENERAL INFORMATION:

APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52

CORRESPONDENCE ADDRESS:

ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-8

Query Match 96.5%; Score 111; DB 3; Length 31;
Best Local Similarity 95.0%; Pred. No. 3.6e-11;
Matches 19: Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TPDINPAMYAGRGIRPVGRF 20
|||||
DB 12 TPDINPAMYTGIRGVGRF 31

RESULT 12

US-09-105-678A-37
Sequence 37, Application US/09105678A
Patent No. 6103882

GENERAL INFORMATION:

APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997

ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-37

Query Match 96.5%; Score 111; DB 3; Length 31;
Best Local Similarity 95.0%; Pred. No. 3,6e-11;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPDINPAMVAGRGIRPVGRF 20
||||| |||||||
DB 12 TPDINPAMVAGRGIRPVGRF 31

RESULT 13
US-09-172-353-4
Sequence 4, Application US/09172353
Patent No. 6197530
GENERAL INFORMATION:
APPLICANT: Stricker-Kongra, Alain
APPLICANT: Gu, Wei
TITLE OF INVENTION: GPRILO AS A TARGET FOR IDENTIFYING WEIGHT MODULATING COMPOUNDS
FILE REFERENCE: 07334/102001
CURRENT APPLICATION NUMBER: US/09/172,353
CURRENT FILING DATE: 1998-10-14
NUMBER OF SEQ ID NOS: 7
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
LENGTH: 31
TYPE: PRT
ORGANISM: Mus musculus
US-09-172-353-4

Query Match 96.5%; Score 111; DB 4; Length 31;
Best Local Similarity 95.0%; Pred. No. 3,6e-11;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPDINPAMVAGRGIRPVGRF 20
||||| |||||||
DB 12 TPDINPAMVAGRGIRPVGRF 31

RESULT 14
US-09-105-678A-38
Sequence 38, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
US-09-105-678A-38

Query Match 96.5%; Score 111; DB 3; Length 32;
Best Local Similarity 95.0%; Pred. No. 3,7e-11;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPDINPAMVAGRGIRPVGRF 20
||||| |||||||
DB 12 TPDINPAMVAGRGIRPVGRF 31

RESULT 15
US-09-105-678A-39
Sequence 39, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440

```

; INFORMATION FOR SEQ ID NO: 39:
;
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 33 amino acids
;   TYPE: amino acid
;   STRANDEDNESS:
;   TOPOLOGY: linear
; MOLECULE TYPE: peptide
;
US-09-105-678A-39

```

```

Query Match          96.5%; Score 111; DB 3; Length 33;
Best Local Similarity 95.0%; Pred. No. 3.9e-11;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 TPDINPAWYAGRGIRPVGRF 20
         |||||  |||||
Db      12 TPDINPAWYTGRIKIRPVGRF 31

```

Search completed: April 17, 2001, 15:39:47
 Job time: 316 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 17, 2001, 15:48:45 ; Search time 39.1 Seconds

Title: US-09-446-543A-8
Perfect score: 115
Sequence: 1 TPDINPAWYAGRGIRPVGRF 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 s
```

Database : Sw1sProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	115	100.0	98	1	PRR_BOVIN	P81264 bos taurus
2	111	96.5	83	1	PRR_RAT	P81278 rattus norv
3	109	94.8	87	1	PRR_HUMAN	P81277 homo sapien
4	106	90.0	424	1	SAMP_SALTY	P23832 salmonella
5	46	40.0	962	1	UVRA_METH	O26543 methanobact
6	45	39.1	798	1	EXL1_HUMAN	O29235 homo sapien
7	45	39.1	798	1	UNR_RAT	P18395 rattus norv
8	44	38.3	973	1	UVRA_RHIME	P56899 rhizobidul m
9	44	38.3	179	1	RK6_GUTTH	O46608 guillardi
10	44	38.3	383	1	CYCB_GYRVI	O82947 chromalum
11	43	37.4	241	1	YGBD_YEAST	P53281 saccharomyc
12	43	37.4	359	1	ALF_HAEN	P44429 haemophilus
13	43	37.4	476	1	YAAU_ECOLI	P30143 escherichia
14	43	37.4	511	1	CP12_CANFA	P56592 canis famli
15	43	37.4	960	1	UVRA_TREPA	O83527 tireponema p
16	43	37.4	1282	1	TP2M_DICDI	P09520 dictyospora
17	42	36.5	332	1	FR2A_DROVI	O02008 drosophila
18	42	36.5	332	1	LYTB_MYCLE	O92781 mycobacteri
19	42	36.5	374	1	VH01_ECOLI	P31393 escherichia
20	42	36.5	453	1	TBR2_GROCN	P32925 geotrichum
21	42	36.5	719	1	ARP_YEAST	P327170 saccharomyc
22	42	36.5	972	1	UVRA_MYCTU	P94972 mycobacteri
23	41.5	36.1	652	1	TEPT_CLOPE	O46506 clostridium
24	41	35.7	264	1	Y355_BUCAT	P57346 buchemera ap
25	41	35.7	313	1	SPEB_STECL	P37819 streptomyce
26	41	35.7	322	1	GRP2_MOUSE	O89100 m grb2-rela
27	41	35.7	330	1	GRP2_HUMAN	O75791 h grb2-rela
28	41	35.7	342	1	Y762_METUA	O58172 methanococc
29	41	35.7	347	1	Y576_METUA	O57996 methanococc
30	41	35.7	376	1	OP51_LIMPO	P35360 limulus pol
31	41	35.7	376	1	OP52_LIMPO	P35361 limulus pol
32	39	35.7	391	1	GAT5_CHICK	P43692 gallus galli
33	41	35.7	423	1	GCRC_MOUSE	P30731 mus musculu

34	4.1	35.7	546	1	CHOD_STRO	P13676 streptomyc
35	4.1	35.7	620	1	YUB8_YEAST	P47069 saccharomy
36	4.1	35.7	652	1	DREB_CHICK	P18302 gallus gall
37	4.1	35.7	775	1	DPOL_THRS9	O56366 thermococ
38	4.1	35.7	943	1	UVRA_HAEIN	O44410 haemophilu
39	4.1	35.7	986	1	CYCR_ARBPV	P11528 arabidopsi
40	4.5	35.2	265	1	UBIE_RICPV	O92cp3 rickettsia
41	40.5	35.2	551	1	ETVA_HUMAN	P43688 homo sapie
42	40.5	35.2	592	1	H1S1_ARATH	O98330 arabidopsi
43	40	34.8	149	1	ENRA_PT77	P00641 bacterioph
44	40	34.8	247	1	MERS_ARATH	P24065 arabidopsi
45	40	34.8	424	1	IMPV_SALTY	P18642 salmonella

ALIGNMENTS

```

RESULT 1
PRR1_BOVIN          STANDARD;      PRT;          98 AA.
ID_PRR1_BOVIN
PR1264;
DT 30-May-2000 (Rel. 39, Created)
DT 30-May-2000 (Rel. 39, Last sequence update)
DT 30-May-2000 (Rel. 39, Last annotation update)
DE PROLACTIN-RELEASING PEPTIDE PRECURSOR (PRR) (PROLACTIN-RELEASING
DE HORMONE) [CONTAINS: PROLACTIN-RELEASING PEPTIDE PRRP1; PROLACTIN-
DE RELEASING PEPTIDE PRRP20].
GN PRR.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
CX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-52.
RP TISSUE=Brain;
RX MEDLINE=98268781; PubMed=9607765;
RA Hituma S., Hebata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,
RA Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,
RA Kurokawa T., Nishimura O., Onda H., Fujino M.;
RT "A prolactin-releasing peptide in the brain.";
RL Nature 393:272-276(1998).
CC -1- FUNCTION: STIMULATES PROLACTIN (PRL) RELEASE AND REGULATES THE
CC EXPRESSION OF PROLACTIN THROUGH ITS RECEPTOR HGR3. MAY STIMULATE
CC LACTOTROPHS DIRECTLY TO SECRETE PRL.
CC FACTOROPHS DIRECTLY TO SECRETE PRL.
CC -1- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC CC EMBL: AB015417; BAA29025.1; -
CC KW Hormone; Amlidation; Signal; Cleavage on pair of basic residues.
CC SIGNAL 1 22 PROLACTIN-RELEASING PEPTIDE PRRP31.
CC PEPTIDE 23 53 PROLACTIN-RELEASING PEPTIDE PRRP20.
CC PEPTIDE 33 53 AMIDATION (G-34 PROVIDE AMIDE GROUP).
CC MOD_RES 53 53
CC SEQUENCE 98 AA: 10544 MW; 08AC35A13H0FA508 CRC64;

```

Query Match	100.0%;	Score 115;	DB 1;	Length 98;
Best Local Similarity	100.0%;	Pred. No. 7.2e-11;		
Matches 20; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

Oy		1 TPDINPAWYAGRCIRPVGRF	20
Db		34 TPDINPAMWZAGRGIRPVGRF	53

RESULT 2
 PRP_RAT STANDARD; PRT: 83 AA.
 AC P81277;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE PROLACTIN-RELEASING PEPTIDE PRECURSOR (PRP) (PROLACTIN-RELEASING
 DE HORMONE) [CONTAINS: PROLACTIN-RELEASING PEPTIDE PRP31; PROLACTIN-
 DE RELEASING PEPTIDE PRP20].
 GN PRH.
 OS Rattus norvegicus (rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=98268781; PubMed=9607765;
 RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,
 RA Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,
 RA Kurokawa T., Nishimura O., Onda H., Fujino M.;
 RA "A prolactin-releasing peptide in the brain."
 RL Nature 393:272-276(1998).
 CC -1- FUNCTION: STIMULATES PROLACTIN (PRL) RELEASE AND REGULATES THE
 CC EXPRESSION OF PROLACTIN THROUGH ITS RECEPTOR HGR3. MAY STIMULATE
 CC LACTOTROPHS DIRECTLY TO SECRETE PRL.
 CC -1- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; AB015418; BAA29026.1;
 KW Hormone; Amidation; Signal; Cleavage on pair of basic residues.
 FT SIGNAL 1 21
 FT PEPTIDE 22 52 PROLACTIN-RELEASING PEPTIDE PRP31.
 FT PEPTIDE 33 52 PROLACTIN-RELEASING PEPTIDE PRP20.
 FT MOD_RES 52 52 AMIDATION (G-53 PROVIDE AMIDE GROUP).
 SQ SEQUENCE 83 AA; 9215 MW; DDC75A264EEB4F29 CRC64;

Query Match 96.5%; Score 111; DB 1; Length 83;
 Best Local Similarity 95.0%; Pred. No. 2.5e-10;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TPDPINPAMYAGRGIRPVGRF 20
 Db 33 TPDPINPAMYAGRGIRPVGRF 52

RESULT 3
 PRP_HUMAN STANDARD; PRT: 87 AA.
 AC P81277;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE PROLACTIN-RELEASING PEPTIDE PRECURSOR (PRP) (PROLACTIN-RELEASING
 DE HORMONE) [CONTAINS: PROLACTIN-RELEASING PEPTIDE PRP31; PROLACTIN-
 DE RELEASING PEPTIDE PRP20].
 GN PRH.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;

RX MEDLINE=98268781; PubMed=9607765;
 RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,
 RA Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,
 RA Kurokawa T., Nishimura O., Onda H., Fujino M.;
 RA "A prolactin-releasing peptide in the brain."
 RL Nature 393:272-276(1998).
 CC -1- FUNCTION: STIMULATES PROLACTIN (PRL) RELEASE AND REGULATES THE
 CC EXPRESSION OF PROLACTIN THROUGH ITS RECEPTOR HGR3. MAY STIMULATE
 CC LACTOTROPHS DIRECTLY TO SECRETE PRL.
 CC -1- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; AB015419; BAA29027.1;
 KW HM; 602663;
 KW Hormone; Amidation; Signal.
 FT SIGNAL 1 22
 FT PEPTIDE 23 53 PROLACTIN-RELEASING PEPTIDE PRP31.
 FT PEPTIDE 34 53 PROLACTIN-RELEASING PEPTIDE PRP20.
 FT MOD_RES 53 53 AMIDATION (G-54 PROVIDE AMIDE GROUP).
 SQ SEQUENCE 87 AA; 9639 MW; 229A2F3F50C981B CRC64;

Query Match 94.8%; Score 109; DB 1; Length 87;
 Best Local Similarity 95.0%; Pred. No. 5.3e-10;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TPDPINPAMYAGRGIRPVGRF 20
 Db 34 TPDPINPAMYAGRGIRPVGRF 53

RESULT 4
 SAMB_SALTY STANDARD; PRT: 424 AA.
 AC P23832;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE SAMB PROTEIN.
 GN SAMB.
 OS Salmonella typhimurium.
 OG Plasmid 60-Mda cryptic.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Salmonella.
 OX NCBI_TaxID=602;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LT2;
 RC MEDLINE=91123176; PubMed=1991707;
 RA Nohmi T., Hakura A., Nakai Y., Watanabe M., Murayama S.Y.,
 RA Sofuni T.;
 RA "Salmonella typhimurium has two homologous but different unimuc
 RT 60-megadon cryptic plasmid of S. typhimurium."
 RL J. Bacteriol. 173:1051-1063(1991).
 CC -1- FUNCTION: INVOLVED IN UV PROTECTION AND MUTATION.
 CC -1- SIMILARITY: BELONGS TO THE IMPB/MCB/SAMB FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----

DR EMBL: D90202; BAAL4226.1; -
 DR PIR: B38176; B38176.
 DR InterPro: IPR001126; -
 DR Pfam: PF00817; IMS: 1.
 KW Plasmid: SOS mutagenesis; DNA repair.
 SQ SEQUENCE 424 AA; 47727 MW; EF8C47476C58A2B CRC64;

Query Match 40.0%; Score 46; DB 1; Length 424;
 Best Local Similarity 53.3%; Pred. No. 9.4;
 Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 2 PDINPAMYAGRGIRP 16
 1:|||||1
 DB 384 PGKGTWFAAGRIAP 398

RESULT 5
 ID VUVA_METHH STANDARD: PRT; 962 AA.
 AC 026543;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE EXCISE-NUCLEASE FROM SUBUNIT A.
 GN VUVA OR MTH43.
 OS Methanobacterium thermoautotrophicum.
 OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
 OC Methanobacter.
 NCBI_TaxID=145262;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-DELTA H;
 RA MEDLINE=98037514; PubMed=9371463;
 RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
 RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
 RA Harrison D., Hoang L., Keagle P., Lum W., Pothier B., Qiu D.,
 RA Spadofora R., Viore R., Wang Y., Wierzbowski J., Gibson R.,
 RA Udwani N., Caruso A., Bush D., Safer H., Patwell J., Prahakar S.,
 RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
 RA Daniels C.J., Mao J.-T., Rice P., Nolling J., Reeve J.N.;
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum
 RT deltar: functional analysis and comparative genomics."
 RT J. Bacteriol. 179:7135-7155(1997).
 CC -1- FUNCTION: THE ABC EXCISION NUCLEASE IS A DNA REPAIR ENZYME THAT
 CC CATALYZES THE EXCISION REACTION OF UV-DAMAGED NUCLEOTIDE SEGMENTS
 CC PRODUCING OLIGOMERS HAVING THE MODIFIED BASE(S). VUVA IS AN ATPASE
 CC AND A DNA-BINDING PROTEIN THAT PREFERENTIALLY BINDS SINGLE-
 CC STRANDED OR UV-IRRADIATED DOUBLE-STRANDED DNA (BY SIMILARITY).
 CC -1- SUBUNIT: CONSISTS OF THREE SUBUNITS: VUVA, VUVB AND VUVC.
 CC -1- SUBCELLULAR LOCATION: CYTOSOLASMIC (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
 CC (ABC TRANSPORTERS). CONTAINS TWO ABC DOMAINS.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AE000828; AAB84949.1; -
 CC InterPro: IPR001617; -
 DR Pfam: PF00005; ABC_tran; 2.
 DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
 KW SOS response; Excision nuclease; DNA repair; ATP-binding; Repeat;
 KW DNA-binding; Zinc-finger.
 FT NP_BIND 38 45
 FT NP_BIND 649 656 ATP (POTENTIAL).
 FT ZN_FING 748 774 C4-TYPE.
 FT SEQUENCE 962 AA; 108395 MW; 2C0EF7C41CCDD060 CRC64;

Query Match 40.0%; Score 46; DB 1; Length 962;
 Best Local Similarity 37.8%; Pred. No. 21;
 Matches 14; Conservative 0; Mismatches 5; Indels 18; Gaps 2;

OY 1 TPINPAMYAG-----RGIRPYGRF 20
 1:|||||1
 DB 704 TPSPNPAITGTGVTHIRLEFAOTPEARKRGYRP-GRF 739

RESULT 6
 ID EXL1_HUMAN STANDARD: PRT; 676 AA.
 AC 029235;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE EXOSTOSIN-LIKE 1 (EXOSTOSIN-L) (MULTIPLE EXOSTOSIS-LIKE PROTEIN).
 GN EXL1 OR EXT1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=97189339; PubMed=9037597;
 RA Wise C.A., Cline G.A., Massa H., Trask B.J., Lovett M.;
 RT "Identification and localization of the gene for EXT1, a third member
 RT of the multiple exostoses gene family."
 RT Genome Res. 7:10-16(1997).
 RL [2]
 RP SEQUENCE FROM N.A.
 RA Xu L., Deng H.X., Xia J.H., Pan Q., Liu C.Y.;
 RT "Mutations of the EXT genes in hereditary multiple exostoses in
 RT Chinese."
 RT Submitted (AUG-1998) to the EMBL/Genbank/DBJ databases.

CC [3]
 RP SEQUENCE FROM N.A.
 RA Wuyts W., Spiekler N., Van Roy N., De Paeppe A., De Bouille K.,
 RA Williams P.J., Van Hul W., Versteeg R., Speleman F.;
 RT "Refined physical mapping and genomic structure of the EXT1 gene."
 RT Submitted (MAY-1999) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: APPEARS TO BE A TUMOR SUPPRESSOR (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ENDOLYSMIC
 CC RETICULUM (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE EXOSTOSIN FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U67191; AAC51141.1; -
 DR EMBL: AF083633; AAD02840.1; JOINED.
 DR EMBL: AF083623; AAD02840.1; JOINED.
 DR EMBL: AF083624; AAD02840.1; JOINED.
 DR EMBL: AF083625; AAD02840.1; JOINED.
 DR EMBL: AF083626; AAD02840.1; JOINED.
 DR EMBL: AF083627; AAD02840.1; JOINED.
 DR EMBL: AF083628; AAD02840.1; JOINED.
 DR EMBL: AF083629; AAD02840.1; JOINED.
 DR EMBL: AF083630; AAD02840.1; JOINED.
 DR EMBL: AF083631; AAD02840.1; JOINED.
 DR EMBL: AF083632; AAD02840.1; JOINED.
 DR EMBL: AF153980; AAF73172.1; -
 DR EMBL: AF153981; AAF73172.1; JOINED.
 KW MIM: 601738; -
 KW Anti-oncogene; Multiple family; Transmembrane; Signal-anchor.
 FT TRANSMEM 10 30
 FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
 FT (POTENTIAL).

FT CARBOHYD 269 269 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 676 AA: 74673 MW: B5E06A8762E5633 CRC64;

Query Match
Best Local Similarity 39.1%; Score 45; DB 1; Length 676;
Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

OY 1 TPDIINPARYAGIRPVGRF 20
DB 400 SPDDPFYVLOGSREGRF 419

RESULT 7
UNR_RAT STANDARD: PRT: 798 AA.

AC P18395;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE UNR_PROTEIN.
GN UNR.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;

Sequence from N.A.
Tissue-Testis:
RX MEDLINE-90370473; PubMed-2204029;
RA Jeffers M., Paclucci R., Pellicer A.;
RL "Characterization of unr, a gene closely linked to N-ras";
CC Nucleic Acids Res. 18:4891-4899(1990).
CC -1- FUNCTION: RNA-BINDING PROTEIN (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: X52311; CA36549.1; -
DR PIR: S11210; S11210.
DR NSSP: P15277; IMJC.
DR InterPro: IPR002059; -
DR Pfam: PF00313; CSD: 8.
DR PROSITE: PS00352; COLD-SHOCK: 4.
KW RNA-binding; Repeat.
FT DOMAIN 26 87 CSD 1.
FT DOMAIN 136 179 CSD 2 (INCOMPLETE).
FT DOMAIN 186 245 CSD 3.
FT DOMAIN 297 337 CSD 4 (INCOMPLETE).
FT DOMAIN 349 410 CSD 5.
FT DOMAIN 447 507 CSD 6.
FT DOMAIN 519 579 CSD 7.
FT DOMAIN 610 670 CSD 8.
FT DOMAIN 674 735 CSD 9.
SQ SEQUENCE 798 AA: 88894 MW: F484B3FA8B095A4 CRC64;

Query Match
Best Local Similarity 39.1%; Score 45; DB 1; Length 798;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 1 TPDIINPARYAGIRPV 17
DB 589 TEEANPTIYSGKVIKPL 605

RESULT 8

UVRA_RHIME STANDARD: PRT: 973 AA.

AC P56899;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE EXCINUCLEASE ABC SUBUNIT A.
GN UVRA.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;

Sequence from N.A.
RP STRAIN-1021;
RA Galibert F., Capela D., Hubler-Barloy F., Gatine M., Batut J.,
RA Boistard P., Guzy J., Kahn D., Thebaud P., Goffeau A.,
RA Purnelle B., Pohl T., Bothe G., Schneider S., Portetelle D.,
RA Vandenbol M., Pehler A., Becker A., Weidner S.;
RL Submitted (MAY-2000) to the SWISS-PROT data bank.

[2]
Sequence of 1-140 from N.A.
RP STRAIN-2021;
RX MEDLINE-99430868; PubMed-10503543;
RA Tapas A., Barbe J.;
RL "Regulation of divergent transcription from the uvra-ssb promoters in
RL Sinorhizobium meliloti";
CC Mol. Gen. Genet. 262:121-130(1999).
CC -1- FUNCTION: THE ABC EXCISION NUCLEASE IS A DNA REPAIR ENZYME THAT
CC CATALYZES THE EXCISION REACTION OF UV-DAMAGED NUCLEOTIDE SEGMENTS
CC PRODUCING OLIGOMERS HAVING THE MODIFIED BASE(S). UVRA IS AN ATPASE
CC AND A DNA-BINDING PROTEIN THAT PREFERENTIALLY BINDS SINGLE-
CC STRANDED OR UV-IRRADIATED DOUBLE-STRANDED DNA (BY SIMILARITY).
CC -1- SUBUNIT: CONSISTS OF THREE SUBUNITS; UVRA, UVRA AND UVRC.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC (ABC TRANSPORTERS). CONTAINS TWO ABC DOMAINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AF125162; AAF03210.1; -
DR InterPro: IPR001617; -
DR PROSITE: PS00211; ABC_TRANSPORTER; PARTIAL.
KW SOS response; Excision nuclease; DNA repair; ATP-binding; Repeat;
KW DNA-binding; Zinc-finger.
FT NE_BIND 34 41 ATP (POTENTIAL).
FT NE_BIND 662 669 ATP (POTENTIAL).
FT ZN_FING 761 787 C4-TYPE.
FT CONFLICT 19 19 G -> A (IN REF. 2).
FT CONFLICT 67 67 F -> S (IN REF. 2).
SQ SEQUENCE 973 AA: 107191 MW: 3E1A8B14527A47FE CRC64;

Query Match
Best Local Similarity 39.1%; Score 45; DB 1; Length 973;
Matches 14; Conservative 2; Mismatches 3; Indels 18; Gaps 3;

OY 1 TPDIINPARYAGIRPVGRF 20
DB 717 TPRSNDATYTGATPDRDWFAGLPEAKARGYP-GRF 752

RESULT 9
RK6_GUTH STANDARD: PRT: 179 AA.
ID RK6_GUTH
AC 046908;

RA van Dyck L., Skala J., de Wergifosse P., Purnelle B., Talla E.,
 RA Nawrocki A., del Bino S., Goffeau A.,
 RL Submitted (May-1996) to the EMBL/Genbank/DBJ databases.
 CC -1- SIMILARITY: CONTRAINS 1 SH3 DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Z72921; CA97149.1; -
 DR HSSP: P29355; ISFM.
 DR SGD: S0003368; YGR136W.
 DR InterPro: IPR001452; -
 DR Pfam: PF00018; SH3; 1.
 DR PRINTS: PR00452; SH3DOMAIN.
 DR PROSITE: PS00002; SH3; 1.
 DR Hypothetical protein; SH3 domain.
 FT DOMAIN 53 112
 SO SEQUENCE 241 AA; 26139 MW; 5F0B1361AF84AA79 CRC64;

Query Match 37.4%; Score 43; DB 1; Length 241;
 Best Local Similarity 41.2%; Pred. No. 16;
 Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 4 INPARYAGRGIRPVGRF 20
 DB 86 ISPMYRGKSNKIKGIF 102

RESULT 12

ALF_HAEIN STANDARD; PRT; 359 AA.

AC P44429;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, last sequence update)
 DT 30-MAY-2000 (Rel. 39, last annotation update)
 DE FRUCTOSE-BISPHOSPHATE ALDOLASE (EC 4.1.2.13).
 GN FBA OR HI0524.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RD / KW20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-T., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Utechtback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhmann J.L., Geoghegan N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus
 RT influenzae Rd.";
 RL Science 269:496-512(1995).
 CC -1- CATALYTIC ACTIVITY: D-FRUCTOSE 1,6-BISPHOSPHATE - GLYCERONE-
 CC PHOSPHATE + D-GLYCERALDEHYDE 3-PHOSPHATE.
 CC -1- COFACTOR: ZINC (BY SIMILARITY).
 CC -1- PATHWAY: SIXTH STEP IN GLYCOLYSIS.
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO CLASS II FRUCTOSE-BISPHOSPHATE ALDOLASE
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U3734; AAC2182.1; -
 DR HSSP: P11604; IDOS.
 DR TIGR: H10524; -
 DR InterPro: IPR000771; -
 DR Pfam: PF01116; F_LBP_aldolase; 1.
 DR PROSITE: PS00806; ALDOLASE_CLASS_II_2; 1.
 DR PROSITE: PS00806; ALDOLASE_CLASS_II_2; 1.
 KM Lyase; Glycolysis; Zinc.
 FT METAL 108 108 ZINC (BY SIMILARITY).
 FT METAL 111 111 ZINC (BY SIMILARITY).
 SO SEQUENCE 359 AA; 39339 MW; 1EDDFCD0969E32C CRC64;

Query Match 37.4%; Score 43; DB 1; Length 359;
 Best Local Similarity 70.0%; Pred. No. 23;
 Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 7 AWYAGRGIRP 16
 DB 67 AFTAGKGRP 76

RESULT 13

YAAI_ECOLI STANDARD; PRT; 476 AA.

AC P30143;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, last sequence update)
 DT 01-NOV-1997 (Rel. 35, last annotation update)
 DE HYPOTHETICAL 51.7 KDA PROTEIN IN THRC-TALB INTERGENIC REGION (ORF8).
 GN YAAU.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=92334977; PubMed=1630901;
 RA Yura T., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N.,
 RA Isono K., Mizobuchi K., Nakata A.;
 RT "Systematic sequencing of the Escherichia coli genome: analysis of
 RT the 0-2.4 min region.";
 RL Nucleic Acids Res. 20:3305-3308(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.M., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
 CC (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE SODIUM:ALANINE SYMPORTER FAMILY
 CC (SAF). STRONG, TO H. INFLUENZAE HI0183.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: D10483; -, NOT_ANNOTATED_CDS.
 DR EMBL: AE000111; AAC73118.1; -.
 DR Ecocore: EG11555; yaad.
 DR InterPro: IPR001463; -.
 DR Pfam: PF01235; Na_Ala_symp; 1.
 DR PRINTS: PR00175; NALASYPOT.
 DR PROSITE: PS00873; NA_ALANINE_SYM; 1.
 KW Hypothetical protein; Transmembrane; Inner membrane; Transport;
 Symport.
 FT TRANSMEM 4 24 POTENTIAL.
 FT TRANSMEM 81 101 POTENTIAL.
 FT TRANSMEM 141 161 POTENTIAL.
 FT TRANSMEM 174 194 POTENTIAL.
 FT TRANSMEM 207 227 POTENTIAL.
 FT TRANSMEM 233 253 POTENTIAL.
 FT TRANSMEM 300 320 POTENTIAL.
 FT TRANSMEM 351 371 POTENTIAL.
 FT TRANSMEM 391 411 POTENTIAL.
 FT TRANSMEM 414 434 POTENTIAL.
 FT SEQUENCE 476 AA; 51662 MW; 216E2E12E126E63 CRC64;
 SO

Query Match 37.4%; Score 43; DB 1; Length 476;
 Best Local Similarity 44.4%; Pred. No. 30;
 Matches 8; Conservative 2; Mismatches 2; Indels 6; Gaps 1;

Qy 3 DIN-----PANTAGRIN 14
 1:1 1111 11:
 Db 120 DVNGGPGGPGAWMARGL 137

RESULT 14
 ID CP12_CANFA STANDARD; PRT; 511 AA.
 AC P56592;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE CYTOCHROME P450 1A2 (EC 1.14.14.1) (CYP1A2) (DMH2) (CYTOCHROME P450-
 DE D2).
 GN CYP1A2.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_Taxid=9615;
 RN [1]
 RP SEQUENCE OF 9-511 FROM N.A.
 RC STRAIN=BEAGLE; TISSUE=Liver;
 RX MEDLINE=91042464; PubMed=2122230;
 RA Uchida T., Komori M., Kitada M., Kametaki T.;
 RT "Isolation of cDNAs coding for three different forms of liver
 RT microsomal cytochrome P-450 from polychlorinated biphenyl-treated
 RT beagle dogs.";
 RT Mol. Pharmacol. 38:644-651(1990).
 RN [2]
 RP SEQUENCE OF 1-16.
 RC STRAIN=BEAGLE; TISSUE=Liver;
 RX MEDLINE=89087526; PubMed=2910310;
 RA Ohno K., Motoya M., Komori M., Miura T., Kitada M., Kametaki T.;
 RT "A novel form of cytochrome P-450 in beagle dogs. P-450-D3 is a low
 RT spin form of cytochrome P-450 but with catalytic and structural
 RT properties similar to P-450d.";
 RT Biochem. Pharmacol. 38:91-96(1989).
 CC -1- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE
 CC MONOOXYGENASES. IN LIVER MICROSOMES, THIS ENZYME IS INVOLVED IN AN
 CC NADPH-DEPENDENT ELECTRON TRANSPORT PATHWAY. IT OXIDIZES A VARIETY
 CC OF STRUCTURALLY UNRELATED COMPOUNDS, INCLUDING STEROIDS, FATTY
 CC ACIDS, AND XENOBIOTICS.
 CC -1- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) = ROH +
 CC OXIDIZED FLAVOPROTEIN + H(2)O.
 CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND. ENDOPLASMIC RETICULUM.
 CC -1- TISSUE SPECIFICITY: CONSTITUTIVELY EXPRESSED IN LIVER.
 CC -1- INDUCTION: BY POLYCHLORINATED BIPHENYL (PCB) IN LIVER AND KIDNEY.

CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR InterPro: IPR001128; -.
 DR InterPro: IPR002401; -.
 DR Pfam: PF00067; P450; 1.
 DR PRINTS: PR00385; P450.
 DR PRINTS: PR00463; EP4501.
 DR PROSITE: PS00086; CYTOCHROME_P450; 1.
 KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
 KW Microsome; Endoplasmic reticulum.
 FT INIT_MET 0
 FT BINDING 453 453 HEME (BY SIMILARITY).
 FT SEQUENCE 511 AA; 57505 MW; 200904C54F4B3CE7 CRC64;
 SO

Query Match 37.4%; Score 43; DB 1; Length 511;
 Best Local Similarity 58.3%; Pred. No. 32;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Oy 1 TPDIINPANTAGR 12
 1:1 1111 11:
 Db 121 SPDSGPVMAAGR 132

RESULT 15
 ID UVRA_TREPA STANDARD; PRT; 960 AA.
 AC O83527;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE EXCINUCLEASE ABC SUBUNIT A.
 GN UVRA OR TP0514.
 OS Treponema pallidum.
 OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
 OX NCBI_Taxid=160;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NICHOLS;
 RX MEDLINE=98332770; PubMed=9665876;
 RA Fraser C.M., Norris S.J., Welstock G.M., White O., Sutton G.G.,
 RA Dodson R., Gwin M., Hickey E.K., Clayton R., Ketchum K.A.,
 RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
 RA Khalak H., Richardson P., Bowman C., Cotton M.D., Fujii C., Garland S.,
 RA McDonald L., Artlich P., Brown J., Cotton M.D., Fujii C., Garland S.,
 RA Hatch B., Horst K., Roberts K., Sandusky N., Weidman J., Smith H.O.,
 RA Venter J.C.;
 RT "Complete genome sequence of Treponema pallidum, the syphilis
 RT spirochete.";
 RT Science 281:375-388(1998).
 CC -1- FUNCTION: THE ABC EXCISION NUCLEASE IS A DNA REPAIR ENZYME THAT
 CC CATALYZES THE EXCISION REACTION OF UV-DAMAGED NUCLEOTIDE SEGMENTS
 CC PRODUCING OLIGOMERS HAVING THE MODIFIED BASE(S). UVRA IS AN ATPASE
 CC AND A DNA-BINDING PROTEIN THAT PREFERENTIALLY BINDS SINGLE-
 CC STRANDED OR UV-IRRADIATED DOUBLE-STRANDED DNA (BY SIMILARITY).
 CC -1- SUBUNIT: CONSISTS OF THREE SUBUNITS: UVRA, UVRB AND UVRC.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
 CC (ABC TRANSPORTERS). CONTAINS TWO ABC DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC EMBL: AE001227; AAC65502.1; -.
 DR TIGR: TP0514; -.
 DR InterPro: IPR001617; -.
 DR Pfam: PF00005; ABC_tran; 2.
 DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
 KW SOS response; Excision nuclease; DNA repair; ATP-binding; Repeat;

KN DNA-binding; zinc-finger.
FT NP_BIND 35 42 ATP (POTENTIAL).
FT NP_BIND 657 664 ATP (POTENTIAL).
FT ZN_FING 270 297 C4-TYPE.
FT ZN_FING 756 782 C4-TYPE.
SQ SEQUENCE 960 AA; 106010 MW; 32F78624B19F7ABF CRC64;

Query Match 37.4%; Score 43; DB 1; Length 960;
Best Local Similarity 35.1%; Pred. NO. 59;
Matches 13; Conservative 1; Mismatches 5; Indels 18; Gaps 2;

QY 1 TPDINPAMYAG-----RGIRPVGRF 20
|| ||| ||
DB 712 TPRSNDPTTYGVFTDIRMLFSQVPEAKMKGYKP-GRF 747

Search completed: April 17, 2001, 15:48:47
Job time: 535 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 17, 2001, 15:45:51 ; Search time 70.08 Seconds
(without alignments)
19.613 Million cell updates/sec

Title: US-09-446-543a-8

Perfect score: 115

Sequence: 1 TPDIHPAWYAGRGIRPYGRF 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08

Maximum Match 100%

Listing first 45 summaries

Database:

1: PIR1:
2: PIR2:
3: PIR3:
4: PIR4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	44.3	430	1 B69009	conserved hypotet
2	50	43.5	527	2 T33175	hypothetical prote
3	50	43.5	790	2 T47959	hypothetical prote
4	48.5	42.2	664	2 F83376	conserved hypotet
5	46	40.0	333	2 H82852	hydroxybenzoate oc
6	46	40.0	424	2 B38176	samb protein - Sal
7	46	40.0	962	2 H69157	exclnuclease ABC c
8	45.5	39.6	779	2 T49717	related to BCS1 pr
9	45	39.1	767	2 F71015	hypothetical prote
10	45	39.1	798	2 S11210	hypothetical prote
11	44	38.3	390	2 G82844	probable untr prote
12	44	38.3	1670	2 S71551	cysteine synthase
13	44	38.3	506	2 H83396	DNA-directed DNA p
14	43.5	37.8	506	2 H83142	probable aldehyde
15	43.5	37.8	1501	2 T45623	hypothetical prote
16	43.5	37.8	220	2 C83292	probable glutathio
17	43	37.4	241	1 S64445	hypothetical prote
18	43	37.4	276	2 D70817	hypothetical prote
19	43	37.4	309	2 T2376	hypothetical prote
20	43	37.4	359	2 C64074	senosy protein k1
21	43	37.4	359	2 T30222	probable amino aci
22	43	37.4	476	2 G64720	sun/nucleolar prot
23	43	37.4	503	2 A82193	cytochrome P450 1A
24	43	37.4	511	2 B37222	hypothetical prote
25	43	37.4	548	2 T47548	exclnuclease ABC c
26	43	37.4	960	2 A71315	DNA topoisomerase
27	43	37.4	1282	2 T30577	probable magnesium
28	43	37.4	1292	2 T31462	
29	43	37.4	1292	2 T31462	

30	42	36.5	143	2 T36978	probable transpos
31	42	36.5	284	2 A75117	hypothetical prote
32	42	36.5	311	2 S66500	cytochrome-c oxida
33	42	36.5	375	2 S47704	hypothetical 41.1K
34	42	36.5	428	2 F81694	pyruvate dehydroge
35	42	36.5	453	2 S18597	tubulin beta chain
36	42	36.5	468	2 C83160	nitrite extrusion
37	42	36.5	719	2 S61046	ARPI protein - yea
38	42	36.5	940	2 A82329	exclnuclease ABC c
39	42	36.5	972	2 A70619	hypothetical prote
40	42	36.5	1296	2 T16859	B1496-FL41 protei
41	42	36.1	165	2 S72776	N-acetylglutamate
42	41.5	36.1	345	2 D84012	hypothetical prote
43	41.5	36.1	443	2 T21499	tetracycline resis
44	41.5	36.1	652	1 S41522	elastic titin - hu
45	41.5	36.1	7962	2 I38346	

ALIGNMENTS

RESULT 1

B69009
conserved hypothetical protein MTH1070 - Methanobacterium thermoautotrophicum (strain C:Species: Methanobacterium thermoautotrophicum
C:Date: 29-Jan-1999 #sequence_revision 29-Jan-1999 #text_change 23-Jul-1999
C:Accession: B69009
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Qiu, D.; Spadlora, R.; Vicaire, R.; Wang, Y.; Mierzbowski, J.; Gibson, R.; Jivani, K.; S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fu
A:Reference number: A69000; MUID:98037514
A:Accession: B69009
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-430 <MTH>
A:Cross-references: GB:AE000877; GB:AE000666; MID:g2622157; PIDN:AA85559.1; PID:g262
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH1070
C:Superfamily: conserved hypothetical protein MTH1070

Query Match 44.3%; Score 51; DB 1; Length 430;
Best Local Similarity 69.2%; Pred. No. 3.5;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 3 DIMPWYAGRGIR 15
DB 191 DIMPWYAGRGIR 203

RESULT 2

T33175
hypothetical protein C246.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T33175
R:Grego, J.; Bradshaw, H.; Keppeler, D.
Submitted to the EMBL Data Library, May 1998
A:Description: The sequence of C. elegans cosmid C246.6.
A:Reference number: Z21298
A:Accession: T33175
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-527 <GRE>
A:Cross-references: EMBL:AB07936; PIDN:AA019213.1; GSPDB:GN00023; CESP:C246.6
A:Experimental source: strain Bristol N2; clone C246
C:Genetics:
A:Gene: CESP:C246.6
A:Map position: 5
A:Introns: 20/3; 77/1; 129/2; 208/3; 470/2

Query Match 43.5%; Score 50; DB 2; Length 527;
 Best Local Similarity 66.7%; Pred. No. 6.1;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 PDINPAMYAGRG 13
 DB 370 PNYLSMYAGRG 381

RESULT 3
 747959
 Hypothetical protein F15G16.60 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
 C:Accession: T47959
 R:De Haan, M.; Maarse, A.C.; Grivell, L.A.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Quer
 submitted to the Protein Sequence Database, January 2000
 A:Reference number: 224480
 A:Accession: T47959
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-790 <DEH>
 A:Cross-references: EMBL:AL132959
 A:Experimental source: cultivar Columbia; BAC clone F15G16
 C:Genetics:
 A:Map position: 3
 A:Introns: 39/71; 678/73; 773/72
 A:Note: F15G16.60

Query Match 43.5%; Score 50; DB 2; Length 790;
 Best Local Similarity 47.4%; Pred. No. 9.2;
 Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 PDINPAMYAGRGIRPVGF 20
 DB 366 PPHNPRTGSRGLQPHGMW 384

RESULT 4
 F83376
 conserved hypothetical protein PA2151 [Imported] - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C:Accession: F83376
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warriner, P.; Hickey, M.J.; Br
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapidis, K.; Lim,
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
 A:Reference number: AB2950; M01D:20437337
 A:Accession: F83376
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-664 <STO>
 A:Cross-references: GB:AE004642; GB:AE004091; NID:g9948163; PIDN:AAG05539.1; GSPDB:GN001
 C:Genetics:
 A:Experimental source: strain PA01
 A:Gene: PA2151

Query Match 42.2%; Score 48.5; DB 2; Length 664;
 Best Local Similarity 62.5%; Pred. No. 13;
 Matches 10; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

QY 1 TPDINPAMYAGRGIRP 16
 DB 478 TPDINP-WFLQSGR 492

RESULT 5

H82852
 hydroxybenzoate octaprenyltransferase XF0068 [Imported] - Xylella fastidiosa (strain
 C:Species: Xylella fastidiosa
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 17-Nov-2000
 C:Accession: H82852
 R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq

A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
 A:Reference number: AB2515; M01D:20365717
 A:Note: for a complete list of authors see reference number A59328 below
 A:Accession: H82852
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1333 <SIM>

A:Cross-references: GB:AE003860; GB:AE003849; NID:g9104830; PIDN:AAF82881.1; GSPDB:GN
 A:Experimental source: strain 9a5c
 R:Stimpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Agencio, M.; Alvarenga, R.
 Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carrero, D.M.; Carrer
 as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.F.; Ferreira, A.J.S.
 submitted to Genbank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La
 Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins
 A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa
 M.; Tsuchiko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.
 A:Reference number: A59328
 A:Contents: annotation
 C:Genetics:
 A:Gene: XF0068
 C:Superfamily: 4-hydroxybenzoate octaprenyltransferase

Query Match 40.0%; Score 46; DB 2; Length 333;
 Best Local Similarity 53.3%; Pred. No. 16;
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 INPAMYAGRGIRPVG 18
 DB 54 LDPYKRLARGDRPVG 68

RESULT 6
 B38176
 samB protein - Salmonella typhimurium

C:Species: Salmonella typhimurium
 C:Date: 28-Aug-1992 #sequence_revision 28-Aug-1992 #text_change 20-Jun-2000
 C:Accession: B38176
 R:Nomai, T.; Nakura, A.; Nakai, Y.; Watanabe, M.; Murayama, S.Y.; Sofuni, T.
 J. Bacteriol. 173, 1051-1063, 1991
 A:Title: Salmonella typhimurium has two homologous but different umuD operons: clon1
 A:Reference number: A38176; M01D:91123176
 A:Accession: B38176
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-424 <NOH>
 A:Cross-references: GB:D90202; NID:g217087; PIDN:BAAL4226.1; PID:g217089
 A:Experimental source: strain LT2
 C:Genetics:
 A:Gene: samB

Query Match 40.0%; Score 46; DB 2; Length 424;
 Best Local Similarity 53.3%; Pred. No. 20;
 Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 2 PDINPAMYAGRGIRP 16

DB 384 PGKIMFAGRIAP 398

RESULT 7

H69157
excinuclease ABC chain A - Methanobacterium thermoautotrophicum (strain Delta H)
N:Alternate names: uvra protein
N:Contains: excision endonuclease ABC (EC 3.1.-.-) chain A
C:Species: Methanobacterium thermoautotrophicum
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 12-Nov-1999
C:Accession: H69157
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
Oliv, D.; Spadofora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.;
Ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funcit
A:Reference number: A69000; MUID:98037514
A:Accession: H69157
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-962 <MTB>
A:Cross-references: GB:AE000628; GB:AE000666; NID:g2621504; PIDN:AB84949.1; PID:g262150
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH443
A:Start codon: TTG
C:Superfamily: excinuclease ABC chain A; ATP-binding cassette homology
C:Keywords: ATP; DNA binding; DNA repair; duplication; hydrolase; P-loop
F:38-45/Region: nucleotide-binding motif A (P-loop)
F:632-915/Domain: ATP-binding cassette homology <ABCE>
F:649-656/Region: nucleotide-binding motif A (P-loop).

Query Match 40.0%; Score 46; DB 2; Length 962;
Best Local Similarity 37.8%; Pred. No. 46;
Matches 14; Conservative 0; Mismatches 5; Indels 18; Gaps 2;

OY 1 TPDPNPMYAG-----RCIRPYGR 20
DB 704 TPSPNPMYGVFPHIRELFAQTPEARKRGYR-GRF 739

RESULT 8
T49717
related to BCS1 protein precursor [Imported] - Neurospora crassa
N:Alternate names: protein B23L21.300
C:Species: Neurospora crassa
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C:Accession: T49717
R:Schulte, U.; Allyn, V.; Hobeisel, J.; Brandt, P.; Partmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
A:Reference number: 225022
A:Accession: T49717
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-779 <SCH>
A:Cross-references: EMBL:AL356172; GSPDB:GN00116; NCSP:B23L21.300
A:Experimental source: BAC clone B23L21; strain OR74A
C:Genetics:
A:Gene: NCSP:B23L21.300
A:Map position: 6

Query Match 39.6%; Score 45.5; DB 2; Length 779;
Best Local Similarity 38.8%; Pred. No. 44;
Matches 10; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

OY 1 TPDPNPMYAGRIAP 14
DB 286 TDTLNPATRRNTANRGI 302

RESULT 9

hypothetical protein PH1420 - Pyrococcus horikoshii
F71015
C:Species: Pyrococcus horikoshii
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 20-Jun-2000
C:Accession: F71015
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, E.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Se
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; Kishida, N.; Ogu
DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophil
A:Reference number: A71000; MUID:98344137
A:Accession: F71015
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-284 <KAW>
A:Cross-references: GB:AE000006; NID:g3236133; PIDN:BAA30526.1; PID:g3257843
A:Experimental source: strain 073
A:Note: this accession replaces an interim accession for a sequence replaced by Genba
C:Genetics:
A:Gene: PH1420
C:Superfamily: Pyrococcus horikoshii hypothetical protein PH1420

Query Match 39.1%; Score 45; DB 2; Length 284;
Best Local Similarity 50.0%; Pred. No. 19;
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 2 PDINPMYAGRIAPYGR 19
DB 217 PYTEPTFALRGLEIGR 234

RESULT 10
T21969
hypothetical protein F38E11.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T21969
R:Matthews, P.
submitted to the EMBL Data Library, January 1996
A:Reference number: Z19495
A:Accession: T21969
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-767 <MTL>
A:Cross-references: EMBL:Z68342; PIDN:CAA92775.1; GSPDB:GN00022; CESP:F38E11.7
A:Experimental source: clone F38E11
C:Genetics:
A:Gene: CESP:F38E11.7
A:Map position: 4
A:introns: 50/2; 118/1; 139/2; 189/3; 226/1; 248/1; 287/2; 375/2; 432/3; 465/3; 548/1

Query Match 39.1%; Score 45; DB 2; Length 767;
Best Local Similarity 75.0%; Pred. No. 52;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 PDINPMY 9
DB 747 PDVKNPMY 754

RESULT 11
S11210
probable untr protein - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 05-Nov-1999
C:Accession: S11210
R:Jeffers, W.; Paclucci, R.; Pellicer, A.
Nucleic Acids Res. 18, 4891-4899, 1990
A:Title: Characterization of untr, a gene closely linked to N-ras.
A:Reference number: S11210; MUID:90370473
A:Accession: S11210

Genotype: DNA

Rislove, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbi, K.; L.

Shadman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Laribag, K.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen
A:Reference number: A82950; MUID:20437337
A:Accession: F83142
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-506 <STO>
A:Cross-references: GB:AE004819; GB:AE004091; NID:g9950214; PIDN:AAG07409.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Gene: PA4022
C:Superfamily: aldehyde dehydrogenase (NAD+); aldehyde dehydrogenase homology

Query Match 37.8%; Score 43.5; DB 2; Length 506;
Best Local Similarity 57.9%; Pred. No. 58;
Matches 11; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 1 TPDINPAMYAGRGIRVGR 19
| | | | | : | | | | : | |
Db 436 TRDINRAYMGGRGIRK-AGR 453

Search completed: April 17, 2001, 15:45:53
Job time: 601 sec



OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 RN NCBI_TaxID=9606;
 RP [1]
 RP SEQUENCE FROM N.A.
 RA Laylor S.;
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF035608; CAB55682.1; -
 FT NON_TER 54
 SO SEQUENCE 54 AA; 6110 MW; E2F3C39F7B961A9F CRC64;

Query Match 49.6%; Score 57; DB 4; Length 54;
 Best Local Similarity 56.2%; Pred. No. 0.11;
 Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 1 TPDPNPAWYAGRGIRP 16
 ||:|||||
 DB 18 TPATPTWYAGSGYYP 33

RESULT 3
 ID 060687 PRELIMINARY; PRT; 465 AA.
 AC 060687;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE SUSHI-REPEAT PROTEIN.
 GN SRPUL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kurosawa H., Inkal T., Inaba T., Goi K., Chang K.-S., Sinjo T.,
 RA Rakesstraw K.M., Naeye C.W., Look T.A.;
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF060567; AAC15765.1; -
 DR INTERPRO: IPR000436; -
 DR INTERPRO: IPR001128; -
 DR PFAM: PF00064; sushi; 3.
 DR PROSITE: PS00086; CYTOCHROME P450; UNKNOWN1.
 DR SEQUENCE 465 AA; 52971 MW; 4D752B187FF3E8B8 CRC64;

Query Match 49.6%; Score 57; DB 4; Length 465;
 Best Local Similarity 56.2%; Pred. No. 1.1;
 Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 1 TPDPNPAWYAGRGIRP 16
 ||:|||||
 DB 18 TPATPTWYAGSGYYP 33

RESULT 4
 ID 027142 PRELIMINARY; PRT; 430 AA.
 AC 027142;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE CONSERVED PROTEIN.
 GN MTH1070.
 OS Methanobacterium thermoautotrophicum.
 OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
 OC Methanobacterium.
 OX NCBI_TaxID=2166;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-DELTA H;
 RX MEDLINE=98037514; PubMed=9371463;

RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
 RA Altridge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
 RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
 RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
 RA Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
 RA McQuigill S., Shiner G., Goyal A., Pietrowski S., Church G.M.,
 RA Daniels C.J., Mao J.-I., Rice P., Nolling J., Reeve J.N.,
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum
 RT delta: functional analysis and comparative genomics."
 RL J. Bacteriol. 179:7135-7155(1997).
 DR EMBL; AE000877; AAB85559.1; -
 DR INTERPRO: IPR002510; -
 DR PFAM: PF01523; pmbA; T1D; 1.
 SO SEQUENCE 430 AA; 46062 MW; C4FAE47C11749E6 CRC64;

Query Match 44.3%; Score 51; DB 1; Length 430;
 Best Local Similarity 69.2%; Pred. No. 8;
 Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 3 DPNPAWYAGRGIR 15
 |||||
 DB 191 DPNPAWYAGRGIR 203

RESULT 5
 ID 076383 PRELIMINARY; PRT; 527 AA.
 AC 076383;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE C24G6.6 PROTEIN.
 GN C24G6.6.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 CC Rhabditidae; Pelodermidae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister L., Latelle P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Ritken L., Roopra A., Saunders D., Showkeen R.,
 RA Smaiden N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.,
 RA "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RA elegans."
 RT Nature 368:32-38(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Greco T., Bradshaw H., Keppler D.;
 RT "The sequence of C. elegans cosmid C24G6."
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Waterston R.;
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF067936; AAC19213.1; -
 DR INTERPRO: IPR000205; -
 DR INTERPRO: IPR002937; -
 DR PFAM: PF01593; Amino-oxidase; 1.
 SO SEQUENCE 527 AA; 59805 MW; 9FBB1FB84437C5CB CRC64;

Query Match 43.5%; Score 50; DB 5; Length 527;
 Best Local Similarity 66.7%; Pred. No. 14;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 PDINPAMYARG 13
 DB 370 PNYLSAMYARG 381

RESULT 6

ID 09M371 PRELIMINARY; PRT; 790 AA.

AC 09M371;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)
 DE 01-OCT-2000 (TREMBLrel. 15, last annotation update)
 GN F15616.60.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA De Haan M., Maarse A.C., Grivell E.A., Nemes H.W., Lemcke K.,
 RA Mayer K.F.X., Queller F., Salanoubat M.;
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL132959; CAB71097.1; -
 KW Hypothetical protein.
 SO SEQUENCE 790 AA; 87375 MW; B222724B75690F30 CRC64;

Query Match 43.5%; Score 50; DB 10; Length 790;
 Best Local Similarity 47.4%; Pred. No. 22;
 Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 2 PDINPAMYARGIRPVGRF 20
 DB 366 PPHNPRTYGRGLOPHGRW 384

RESULT 7

ID 074569 PRELIMINARY; PRT; 343 AA.

AC 074569;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, last annotation update)
 DE RAHL.
 GN RAHL.
 OS Coprinus cinereus (Inky cap fungus).
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Agaricales;
 OC Coprinaceae; Coprinus.
 OC NCBI_TaxID=5346;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN-OKAYAMA-7;
 RC MEDLINE=91174112; PubMed=9021132;
 RX Yeager Stassen N., Logsdon J.M. Jr., Vora G.J., Offenberg H.H.,
 RA Palmer J.D., Zolan M.E.;
 RT "Isolation and characterization of rad51 orthologs from Coprinus
 cinereus and Lycopericon esculentum, and phylogenetic analysis of
 eukaryotic recA homologs."
 RT Curr. Genet. 31:144-157(1997).
 RL EMBL; U21905; AAC23703.1; -
 DR EMBL; U21905; AAC23703.1; -
 DR INTERPRO; IPR001553; -
 DR PFAM; PF00633; HHH; 1.

SO SEQUENCE 343 AA; 36789 MW; EA8AA6BB5B83D226 CRC64;

Query Match 41.7%; Score 48; DB 3; Length 343;
 Best Local Similarity 50.0%; Pred. No. 18;
 Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

OY 1 TPDINPAMYARGIRPVG 18
 DB 273 TPDAPGPYAGNEKPKPG 290

RESULT 8

ID 09RJ10 PRELIMINARY; PRT; 333 AA.

AC 09RJ10;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
 DE 01-MAY-2000 (TREMBLrel. 13, last annotation update)
 GN SCF73.06C.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OC NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Seeger K.D., Harris D.;
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA STRAIN-A3(2);
 RC Parthill J., Bartell B.G., Rajandream M.A.;
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA STRAIN-A3(2);
 RC Redenbach M., Kieser H.W., Denapalte D., Eichner A., Cullum J.,
 RA Kinsahl H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 the 8 Mb Streptomyces coelicolor A3(2) chromosome."
 RT Mol. Microbiol. 21:77-96(1996).
 DR EMBL; AL121746; CAB57411.1; -
 KW Hypothetical protein.
 SO SEQUENCE 333 AA; 36312 MW; 884FF2861837FD76 CRC64;

Query Match 40.4%; Score 46.5; DB 2; Length 333;
 Best Local Similarity 35.5%; Pred. No. 29;
 Matches 11; Conservative 2; Mismatches 7; Indels 11; Gaps 1;

OY 1 TPDINPAM-----YAGRGIRPVGRF 20
 DB 19 TPWEPAPWTEATGWTARLAHLRPTGRW 49

RESULT 9

ID 09PH76 PRELIMINARY; PRT; 333 AA.

AC 09PH76;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, last annotation update)
 DE HYDROXYBENZATE OCTAPRENYLTRANSFERASE.
 GN XE0068.
 OS Xylella fastidiosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
 OC Xylella.
 OC NCBI_TaxID=2371;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-9A5C;

RX MEDLINE-20365717; PubMed-10910347;
 RA Simpson A.J.G., Reinech F.C., Arruda P., Abreu F.A., Acencio M.,
 RA Alvaranga R., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S.,
 RA Barros M.N., Bonaccorsi E.D., Bordin S., Bove J.M., Biltons M.R.S.,
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carter N.,
 RA Coluto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 RA Coutinho L.P., Cristofani M., Dias-Neto E., Docena C., El-Dorry N.,
 RA Facchini A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
 RA Fraga J.S., Franca S.C., Franco M.C., Frohne M., Furian L.R.,
 RA Gartner M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 RA Ho P.L., Hohelsel J.D., Junqueira M.L., Kemper E.U., Kitajima J.P.,
 RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 RA Machado M.A., Madela A.M.B.N., Madela H.M.F., Marinho C.L.,
 RA Marques M.V., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 RA Nhani A.J., Nobrega F.G., Nunes L.R., Oliveira M.A., Paris A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Pariz A.,
 RA Pelxoto B.R., Pereira G.A.G., Pereira H.A., Jr., Pesquero J.B.,
 RA Queiroz R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
 RA de Rosa V.E., Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A., Jr.,
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuhako M.H.,
 RA Vallada H., Van Slyks M.A., Verjovski-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Zeldanis J., Zetubal J.C.,
 RT The genome sequence of the plant pathogen *Xylella fastidiosa*.;
 RL Nature 406:151-157(2000).
 DR EMBL; AE003860; AAF82881.1; -
 DR INTERPRO; IPR000537; -
 DR PFAM; PF01040; ubia; 1.
 DR PROSITE; PS00943; UBLA; ECF3F4716C962B95 CRC64;
 SQ SEQUENCE 333 AA; J7931 MW; ECF3F4716C962B95 CRC64;

Query Match 40.0%; Score 46; DB 2; Length 333;
 Best Local Similarity 53.3%; Pred. No. 35;
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 INPAMYAGRGIRPVG 18
 : : : : :
 DB 54 LDPMYKLAGDRPVG 68

RESULT 10
 000050 PRELIMINARY; PRT; 555 AA.

AC 000050;
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
 DE TRANSPONASE.
 OS Aspergillus awamori.
 OC Eukaryota; Fungi; Ascomycota; Eurotiiales; Trichocomaceae;
 OC anamorphic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=105351;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OVK143F DERIVED FROM NRRL3112; TRANSPON-VADEP;
 RA MEDLINE-97156905; PubMed-9003286;
 RA Nyssonen E., Amutan M., Enfield L., Stubbs J., Dunn-Coleman N.S.,
 RT "The transposable element TanI of *Aspergillus niger* var. *awamori*, a
 RT new member of the FcII family.";
 RL Mol. Genet. 253:50-56(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OVK143F DERIVED FROM NRRL3112; TRANSPON-VADEP;
 RA MEDLINE-96207472; PubMed-8655427;
 RA Amutan M., Nyssonen E., Stubbs J., Diaz-Torres M.R., Dunn-Coleman N.,
 RT "Identification and cloning of a mobile transposon from *Aspergillus*
 RT *niger* var. *awamori*.";
 RL Curr. Genet. 29:468-473(1996).

DR EMBL; U58946; AAC49623.1; -
 SQ SEQUENCE 555 AA; 62978 MW; 70904D2EE09EA33 CRC64;

Query Match 40.0%; Score 46; DB 3; Length 555;
 Best Local Similarity 57.1%; Pred. No. 60;
 Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 7 AWYAGRGIRPVGRP 20
 : : : : :
 DB 237 AWYAGRGIRPVGRP 250

RESULT 11
 09NE93 PRELIMINARY; PRT; 820 AA.

AC 09NE93;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE HYPOTHETICAL 86.4 KDA PROTEIN.
 GN L787.05.
 OS Leishmania major.
 OC Eukaryota; Euzoenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 OX NCBI_TaxID=5664;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FRIEDLIN;
 RA Robben J., Aert R., Volckaert G., Ivens A.C., Quail M.,
 RA Rajadream M.A., Barrell B.G.,
 RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FRIEDLIN;
 RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
 RA Smith D.F.,
 RT "A physical map of the *Leishmania major* Friedlin genome.";
 RL Genome Res. 8:135-145(1998).
 DR EMBL; AL163492; CAB86689.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 820 AA; 86388 MW; F4CBDD9FDE66817 CRC64;

Query Match 40.0%; Score 46; DB 5; Length 820;
 Best Local Similarity 47.1%; Pred. No. 91;
 Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 PDINPAMYAGRGIRPVG 18
 : : : : :
 DB 406 PPTSPALYSGNGSOPLG 422

RESULT 12
 09P5J9 PRELIMINARY; PRT; 779 AA.

AC 09P5J9;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE RELATED TO BC51 PROTEIN PRECURSOR.
 GN B23L21.300.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Sordariales; Sordariaceae; Neurospora.
 OX NCBI_TaxID=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Schulte U., Algn V., Nohelsel J., Brandt P., Fartmann B., Holland R.,
 RA Nyakatura G., Mewes H.W., Mannhaupt G.,
 RL Submitted (May-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA German Neurospora genome project;
 RL Submitted (May-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AL356172; CAB91698.1; -.
SQ SEQUENCE 779 AA; 85796 MW; CA7891402DFFBE30 CRC64;

Query Match:	39.6%;	Score 45.5;	DB 3;	Length 779;
Best Local Similarity:	98.8%;	Pred. No. 1e+02;	3;	Indels 3;
Matches 10;	Conservative	1;	Mismatches	3;
Gaps	1;			
Qy	1	TPDINPA---WVAGRGI	14	
Db	286	TDVINDPATRRIRNRGI	302	

RESULT	13	
OS0128		
ID	OS0128	PRELIMINARY;
AC	OS0128;	PRT; 284 AA.
DT	01-JUN-1998 (Tremblrel, 06, Created)	
DT	01-JUN-1998 (Tremblrel, 06, Last sequence update)	
DT	01-JUN-2000 (Tremblrel, 14, Last annotation update)	
DE	HYDROPHETICAL 32.3 KDA PROTEIN PH1420.	
GN	PH1420.	
OS	Pyrococcus horikoshii.	
OC	Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus	
CC	NCBI_TaxId=53953;	
LN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN-OT3;	
RX	MEDLINE=98344137; PubMed=9679194;	
RA	Kawarabayashi Y., Sawada M., Horikawa H., Hino Y.,	
RA	Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.	
RA	Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,	
RA	Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi,	
RA	Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,	
RA	Maaschik Y., Shizuya H., Kikuchi H.;	
RT	"Complete sequence and gene organization of the genome of a hyper-thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";	
RL	DNA Res. 5:55-76(1998).	
DR	EMBL; AP000006; BAA30526.1; -;	
DR	INTERPRO; IPR001066; -;	
DR	PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.	
KW	Hypoethetical protein	
SQ	SEQUENCE 284 AA; 32319 MW; 8E0E7BC3711D3815 CRC64;	

Query Match	39.1%	Score 45;	DB 1;	length 284;
Best local Similarity	50.0%	Pred. No. 42;		
Matches	9;	Conservative	3;	Mismatches
			6;	Indels
			0;	Gaps
			0;	
QY	2	PDINPAMVYAGRGIRPVGR	19	
DB	217	PTLEPTFTALRGLELDLGR	234	

RESULT 14
 Q020170
 ID Q20170 PRELIMINARY; PRT; 767 AA.
 AC Q20170;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
 DE F38E11.7 PROTEIN.
 GN F38E11.7.
 OS *Ceenorhabditis elegans*.
 OC Eukaryota; Metazoa; Nematoda; Rhabditida; Rhabditidae.
 OC Rhabditidae; Palodeirinae; *Ceenorhabditis*.
 OX NCBI_TaxID=6239;
 RN [1]
 RF SEQUENCE FROM N.A.
 RA Matthews P.;
 RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
 RP [2]
 RP SEQUENCE FROM N.A.

RX MEDLINE:94150718; PubMed:9606398;
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berns M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Dublin R., Faveilo A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Keshav J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Riken L., Roopra A., Saunders D., Showkneen R.,
 RA Smalodon N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Wellstock L., Wilkinson-Sproat J., Wohldman P.,
 RT 7.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RL elegans. "-
 RL Nucleic 368:32-38(1994).
 DR EMBL: Z68342; CA92775.1; "-
 DR INTERPRO: IPR000595; "-
 DR INTERPRO: IPR001623; "-
 DR INTERPRO: IPR002025; "-
 DR PFAM: PF00027; CNMP_binding; 1.
 DR PFAM: PF00914; CNG_membrane; 1.
 DR PROSITE: PS00888; CNMP_BINDING_1; UNKNOWN_1.
 DR PROSITE: PS50042; CNMP_BINDING_3; 1.
 SQ SEQUENCE 767 AA; 8998 MW; F7EFC69DBBEACF3 CRC64;

Query Match	39.1%	Score 45;	DB 5;	Length 767;
Best Local Similarity	75.0%	Pred. No. 1.2e+02;		
Matches	6;	Conservative	1;	Mismatches 1;
				Indels 0;
				Gaps 0;
QY	2 PDINPAMY	9		
	:			
Db	747 PDVKPAMY	754		

Q9T133	15	PRELIMINARY;	PRT:	153 AA.
ID	Q9T133			
DC	Q9T133;			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, last sequence update)			
DT	01-MAY-2000 (TREMBLrel. 13, last annotation update)			
DE	ENDONUCLEASE.			
GN	3.			
OS	Bacteriophage phi-yeo3-12.			
OC	Viruses; dsDNA viruses, no RNA stage; Tailed phages; Podoviridae;			
OC	T7-like phages.			
OX	NCBI_TaxId=110457;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Pajunen M.I., Kiljunen S.J., Skurnik M.,			
RT	"Complete genomic sequence of the lytic Bacteriophage phi-yeo3-12 of			
RT	<i>Yersinia enterocolitica</i> serotype O:3.",			
RL	Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.			
FM	EMBL: AJ251805; CAB63604.1; "			
KW	Endonuclease.			
SQ	SEQUENCE 153 AA; 17640 BP; 21157BBDE6C641D CRC64;			

Query Match	38.3%	Score 44	DB 9	length 153
Best Local Similarity	66.7%	Pred. No. 30		
Matches	8	Conservative	1	Mismatches 3
				Indels 0
				Gaps 0
OY	9	YAGRGIRVGRF	20	
	11	1111111	1	
Db	5	YAARGVRKVGAF	16	

search completed: April 17, 2001, 15:48:06
Job time: 559 sec

Tue Apr 17 15:46:29 2001

us-09-446-543a-8.ispt

Page 6

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 17, 2001, 15:45:53 ; Search time 70.08 Seconds
(without alignments)
30.400 Million cell updates/sec

Title: US-09-446-543a-47
Perfect score: 174
Sequence: 1 SRAQHSMTETPDINPAMYTGRCIRPVGRF 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database: PIR 67: *
1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	61	35.1	1236	2 T50904	Mg protoporphyrin
2	58	33.3	1232	2 T31462	probable magnesium
3	56	32.2	962	2 H69157	exci-nuclease ABC c
4	53	30.5	798	2 S11210	probable untr prote
5	53	30.5	1415	2 C83070	conserved hypochet
6	52	29.9	940	2 A82329	exci-nuclease ABC,
7	52	29.9	972	2 A70619	exci-nuclease ABC c
8	51.5	29.6	503	2 A82193	Sun/nucleolar prote
9	51	29.3	294	2 T21075	hypothetical prote
10	51	29.3	943	2 D64057	exci-nuclease ABC c
11	50.5	29.0	176	2 S67150	hypothetical prote
12	50.5	29.0	548	2 T47548	hypothetical prote
13	50	28.7	482	1 S40887	RS167 protein - Y
14	50	28.7	940	1 BYECNA	exci-nuclease ABC c
15	50	28.7	953	2 D71645	exci-nuclease ABC s
16	50	28.7	965	2 C82560	exci-nuclease ABC c
17	49.5	28.4	375	2 S47704	hypothetical 41.1K
18	49	28.2	128	2 S76955	hypothetical prote
19	49	28.2	220	2 C83292	probable glutathio
20	49	28.2	772	2 T07958	protoporphyrin IX
21	49	28.2	943	2 E83117	exci-nuclease ABC s
22	49	28.2	952	1 T46550	exci-nuclease ABC c
23	49	28.2	970	2 S77349	exci-nuclease ABC c
24	49	28.2	1014	2 T36031	magnesium-protopor
25	49	28.2	1193	2 T50729	protoporphyrin IX
26	49	28.2	1331	2 S75000	protoporphyrin IX
27	49	28.2	1379	2 S37310	protoporphyrin IX
28	49	28.2	1380	2 S64721	protoporphyrin IX
29	49	28.2	1381	2 S71288	protoporphyrin IX

30	49	28.2	1382	2 T01789	protoporphyrin IX
31	49	28.2	1383	2 T07126	magnesium chelatase
32	49	28.2	1670	2 S71551	DNA-directed DNA p
33	48.5	27.9	664	2 F83376	conserved hypochet
34	48	27.6	157	2 A81611	hypothetical prote
35	48	27.6	455	2 D70885	probable aldC prot
36	48	27.6	498	2 T09022	beta-glucosidase h
37	48	27.6	517	2 T09022	ARPI protein - yea
38	48	27.6	719	2 S61046	hypothetical prote
39	48	27.6	790	2 T47959	N ras upstream pro
40	48	27.6	798	2 S29815	exci-nuclease ABC c
41	48	27.6	948	2 B81883	exci-nuclease ABC c
42	48	27.6	949	2 A81138	exci-nuclease ABC c
43	48	27.6	960	2 A71315	magnesium-protopor
44	48	27.6	1194	2 D49851	TMF receptor assoc
45	47.5	27.3	501	2 T61512	

ALIGNMENTS

RESULT 1
T50904
Mg protoporphyrin methyl transferase [imported] - Rubrivivax gelatinosus
C:Species: Rubrivivax gelatinosus
C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 02-Sep-2000
C:Accession: T50904
R:Nagashima, K.V.; Igarashi, N.; Harada, J.; Nagashima, S.; Matsuura, K.; Shimada, K.
A:Description: Determination of Nucleotide Sequences of Rubrivivax gelatinosus Photos
submitted to the EMBL Data Library, November 1999
A:Reference number: Z25270
A:Accession: T50904
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1236 <NAG>
A:Cross-references: EMBL:AB034704; PDB:BAA94057.1
A:Experimental source: strain T1144
C:Genetics: bchH
A:Gene: bchH
C:Superfamily: Rhodobacter capsulatus magnesium-protoporphyrin O-methyltransferase

Query Match 35.1%; Score 61; DB 2; Length 1236;
Best Local Similarity 37.5%; Pred. No. 1.3;
Matches 12; Conservative 6; Mismatches 8; Indels 6; Gaps 1;
QY 3 AHOHSMETRPDINPAMYTG-----RGIRPV 28
DB 1112 SEQVALETETRLNPKYEGMELHGEYGRKI 1143

RESULT 2
T31462
probable magnesium chelatase (EC 4.99.1.-) chain H bchH - Hellobacillus mobilis
C:Species: Hellobacillus mobilis
C:Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 15-Sep-2000
C:Accession: T31462
R:Xiong, J.; Inoue, K.; Bauer, C.E.
A:Title: Tracking molecular evolution of photosynthesis by characterization of a major
A:Reference number: Z21036; M01D:99061957
A:Accession: T31462
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1292 <XIO>
A:Cross-references: EMBL:AF080002; M1D:93820536; P1D:93820560; P1D:AA084033.1
C:Genetics: bchH
A:Gene: bchH
C:Superfamily: Rhodobacter capsulatus magnesium-protoporphyrin O-methyltransferase
C:Keywords: lyase

Query Match 33.3%; Score 58; DB 2; Length 1292;

Best Local Similarity 42.3%; Pred. No. 3.8;
Matches 11; Conservative 3; Mismatches 6; Indels 6; Gaps 1;

9 ETRTPDINPAMWTG-----KQIRPV 28

Db 1176 ETRITLANKMKGKMGKGVREI 1201

RESULT 3
H69157

excinuclease ABC chain A - Methanobacterium thermoautotrophicum (strain Delta H)

N:Contains: excision endonuclease ABC (EC 3.1.-.-) chain A

C:Species: Methanobacterium thermoautotrophicum

C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 12-Nov-1999

C:Accession: H69157

R:Smith, D.R.; Doucette-Stamm, L.A.; Delonguey, C.; Lee, H.; Dubois, J.; Aldredge, T.;

Qi, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.

J. Bacteriol. 179, 7135-7155, 1997

A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct

A:Reference number: A69000; MUID:98037514

A:Accession: H69157

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Residues: 1-962 <MTH>

A:Cross-references: GB:AE000828; GB:AE000666; NID:92621504; PIDN:AA84949.1; PID:9262150

A:Experimental source: strain Delta H

C:Genetics:

A:Gene: MTH443

A:Start codon: TTG

C:Superfamily: excinuclease ABC chain A; ATP-binding cassette homology

C:Keywords: ATP; DNA binding; DNA repair; duplication; hydrolase; P-loop

F:38-45/Region: nucleotide-binding motif A (P-loop)

F:632-915/Domain: ATP-binding cassette homology <ABCE>

F:649-656/Region: nucleotide-binding motif A (P-loop)

Query Match 32.2%; Score 56; DB 2; Length 962;

Best Local Similarity 42.1%; Pred. No. 5.4;

Matches 16; Conservative 0; Mismatches 4; Indels 18; Gaps 2;

11 RTPDINPAMWTG-----KQIRPVGRF 31

Db 703 RTPRNPATYTGFTPIRELPACTPARRKGRYP-GRF 739

RESULT 4

S11210

probable unr protein - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 05-Nov-1999

C:Accession: S11210

R:Jeffers, M.; Paculci, R.; Pellicer, A.

Nucleic Acids Res. 18, 4891-4899, 1990

A:Title: Characterization of unr: a gene closely linked to N-ras.

A:Reference number: S11210; MUID:90370473

A:Accession: S11210

A:Molecule type: mRNA

A:Residues: 1-798 <JEF>

A:Cross-references: EMBL:552311; NID:957454; PIDN:CAA36549.1; PID:957455

C:Keywords: DNA binding

Query Match 30.5%; Score 53; DB 2; Length 798;

Best Local Similarity 43.5%; Pred. No. 12;

Matches 10; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

6 HSMETRPDINPAMWTGRCIRPV 28

Db 583 HSYNGITEPANTPIYSKVIKRLP 605

RESULT 5

C83070

conserved hypothetical protein PA4601 [Imported] - Pseudomonas aeruginosa (strain PAO

C:Species: Pseudomonas aeruginosa

C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C:Accession: C83070

R:Stover, C.K.; Pham, X.O.; Ewlin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;

Adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; L

.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa

A:Reference number: AB2950; MUID:20437337

A:Accession: C83070

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1415 <STO>

A:Cross-references: GB:AE004874; GB:AE004091; NID:99950849; PIDN:AA07989.1; GSPDB:GN

A:Experimental source: strain PA01

C:Genetics:

A:Gene: PA4601

Query Match 30.5%; Score 53; DB 2; Length 1415;

Best Local Similarity 37.5%; Pred. No. 22;

Matches 9; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

1 SRANSHMETRPDINPAMWTGRC 24

Db 330 AQARQHLLGLLPDLPGMLTDAG 353

RESULT 6

AB2329

excinuclease ABC, chain A VC0394 [Imported] - Vibrio cholerae (group O1 strain N1661

C:Species: Vibrio cholerae

C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 15-Sep-2000

C:Accession: AB2329

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.

chardson, D.; Emolae, M.D.; Yamahyan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers

L, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: AB2035; MUID:20406833

A:Accession: AB2329

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-940 <HEI>

A:Cross-references: GB:AE004127; GB:AE003852; NID:9654808; PIDN:AAF93567.1; GSPDB:GN

A:Experimental source: serogroup O1; strain N1661; biotype El Tor

C:Genetics:

A:Gene: VC0394

A:Map position: 1

C:Superfamily: excinuclease ABC chain A; ATP-binding cassette homology

Query Match 29.9%; Score 52; DB 2; Length 940;

Best Local Similarity 39.5%; Pred. No. 20;

Matches 15; Conservative 1; Mismatches 4; Indels 18; Gaps 2;

11 RTPDINPAMWTG-----KQIRPVGRF 31

Db 694 RTPRNPATYTGFTPIRELPACTPARRKGRYP-GRF 730

RESULT 7

A70619

excinuclease ABC chain A - Mycobacterium tuberculosis (strain H37RV)

N:Contains: excision endonuclease ABC (EC 3.1.-.-) chain A

C:Species: Mycobacterium tuberculosis

C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000

C:Accession: A70619

R:cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon

Connor, R.; Davies, R.; Devlin, K.; Fellwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrall, B.G.
A>Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
A:Reference number: A70500; M01D:98293987
A:Accession: A70619
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-972 <COL>
A:Cross-references: GB:Z65982; GB:AL123456; M1D:93261718; P1D:CA006633.1; P1D:91838989
A:Experimental source: strain H37Rv
C:Genetics:
A:Gene: *uvrA*
C:Superfamily: excinuclease ABC chain A; ATP-binding cassette homology
C:Keywords: ATP; DNA binding; DNA repair; hydrolase
F:32-38/Region: nucleotide-binding motif A (P-loop)
F:537-920/Domain: ATP-binding cassette homology <ABCE>
F:654-661/Region: nucleotide-binding motif A (P-loop)

Query Match 29.9%; Score 52; DB 2; Length 972;
Best Local Similarity 39.5%; Pred. No. 21;
Matches 15; Conservative 1; Mismatches 4; Indels 18; Gaps 2;

QY 11 RTPDINPWTG-----RGIRPGRP 31
DB 708 RTPRSNPRTYGVFKIRFLFATTEAKRGYGP-GRF 744

RESULT 8
Sun/nucleolar protein family protein VCI502 [Imported] - Vibrio cholerae (group O1 strain A82193)
C:Species: *Vibrio cholerae*
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 15-Sep-2000
C:Accession: A82193
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.; Chardson, D.; Ermolova, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, F. L.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A>Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
A:Reference number: A82035; M01D:20406833
A:Accession: A82193
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-503 <HEB1>
A:Cross-references: GB:AE004228; GB:AE003852; M1D:9655997; P1D:AA94657.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VCI502
A:Map position: 1

Query Match 29.6%; Score 51.5; DB 2; Length 503;
Best Local Similarity 36.7%; Pred. No. 12;
Matches 11; Conservative 5; Mismatches 13; Indels 1; Gaps 1;

QY 1 SRAHQSHETRTPDINPWTGKIRPGVR 30
DB 435 SSSASHVEIDTQAR-EMFGKCVREPGQ 463

RESULT 9
T21075
Hypothetical protein F19H6.1 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000
C:Accession: T21075; T21124
R:McMurray, A.
Submitted to the EMBL Data Library, August 1995
A:Reference number: Z19368
A:Accession: T21075
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
A:Residues: 1-294 <W12>
A:Cross-references: EMBL:Z50873; P1D:CAA90762.1; GSPDB:GN00028; CESP:F19H6.1
A:Experimental source: clone F17E5
R:McMurray, A.
Submitted to the EMBL Data Library, November 1995
A:Reference number: Z19378
A:Accession: T21124
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-294 <W12>
A:Cross-references: EMBL:Z68115; P1D:CAA92169.1; GSPDB:GN00028; CESP:F19H6.1
A:Experimental source: clone F19H6
C:Genetics:
A:Gene: CESP:F19H6.1
A:Map position: X:85/2; 113/3; 150/1; 192/3; 266/2
A:Introns: 55/3; 85/3; 85/3
C:Superfamily: kinase-related transforming protein; protein kinase homology

Query Match 29.3%; Score 51; DB 2; Length 294;
Best Local Similarity 44.8%; Pred. No. 7.9;
Matches 13; Conservative 2; Mismatches 12; Indels 2; Gaps 1;

QY 3 AHQSHETRTPDINPA--WTGKIRPGV 29
DB 139 AHHSKRIMRDIKPAWVFYTGNGIKLG 167

RESULT 10
D64057
exinuclease ABC chain A - *Haemophilus influenzae*
N:Alternate names: *uvrA* protein
C:Species: *Haemophilus influenzae*
C>Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 12-Nov-1999
C:Accession: D64057; J05157
R:Feilschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage
Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.; Glodek, A.; Kelley, J.M.; Weidman
D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A:Authors: Guehen, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter
A>Title: Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd.
A:Reference number: A64000; M01D:95350650
A:Accession: D64057
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-943 <RIGR>
A:Cross-references: GB:U32711; GB:U42023; M1D:91573214; P1D:AA021915.1; P1D:91573215
A:Experimental source: strain Rd KW20
R:de la Morena, M.L.; Hendrixson, D.R.; St Gene III, J.W.
Gene 177, 23-28, 1996
A>Title: Isolation and characterization of the *Haemophilus influenzae* *uvrA* gene.
A:Reference number: J05157; M01D:97080495
A:Accession: J05157
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-162; 'L', 164-235; 'D', 237-424; 'K', 426-462; 'W', 464-513; 'Q', 515-660; 'T', 662
A:Cross-references: GB:U33877; M1D:91144488; P1D:AA044592.1; P1D:91144489
A:Experimental source: strain N187
C:Genetics:
A:Gene: *uvrA*
C:Function:
A:Description: has ATPase and DNA binding activity; involved in DNA repair
C:Superfamily: excinuclease ABC chain A; ATP-binding cassette homology
C:Keywords: ATP; DNA binding; DNA repair; duplication; hydrolase; P-loop
F:31-38/Region: nucleotide-binding motif A (P-loop)
F:623-907/Domain: ATP-binding cassette homology <ABCE>
F:640-647/Region: nucleotide-binding motif A (P-loop)

Query Match 29.3%; Score 51; DB 2; Length 943;
Best Local Similarity 39.5%; Pred. No. 28;

C:Comment: This protein is an ATPase and a DNA-binding protein that preferentially binds e, an ATP-dependent DNA repair enzyme that catalyzes the excision reaction of UV-damaged C:Genetics:

A:Gene: uvra

A:Map position: 92 min

C:Superfamily: excinuclease ABC chain A; ATP-binding cassette homology

C:Keywords: ATP; DNA binding; DNA repair; duplication; hydrolase; P-loop

F:3138/Region: nucleotide-binding motif A (P-loop)

F:623-907/Domain: ATP-binding cassette homology <ABCE>

F:640-647/Region: nucleotide-binding motif A (P-loop)

Query Match

Best Local Similarity 39.5%; Score 50; DB 1; Length 940;
Matches 15; Conservative 0; Mismatches 5; Indels 18; Gaps 2;

QY 11 RPPDINPAMWTG-----RGIRPVGRF 31

DB 695 RPPRSNPATYTGVEFVRELFGVPSRARGYTP-CRF 731

RESULT 15

D71645

excinuclease ABC chain A (uvra) RP835 - Rickettsia prowazekii

N:Contains: excision endonuclease ABC (EC 3.1.-.-) chain A

C:Species: Rickettsia prowazekii

C>Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 03-Nov-2000

C:Accession: D71645

R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Slicheritz-Ponten, T.; Alsmark, D

Nature 396, 133-140, 1998

A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.

A:Reference number: A71630; MUID:99039499

A:Accession: D71645

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-953 <AND>

A:Cross-references: GB:AJ235273; GB:AJ235269; NID:g3861237; PID:CA15260.1; PID:g386136

A:Experimental source: strain Madrid E

C:Genetics:

A:Gene: uvra; RP835

C:Superfamily: excinuclease ABC chain A; ATP-binding cassette homology

C:Keywords: ATP; DNA binding; DNA repair; hydrolase

F:33-40/Region: nucleotide-binding motif A (P-loop)

F:635-919/Domain: ATP-binding cassette homology <ABCE>

F:652-659/Region: nucleotide-binding motif A (P-loop)

Query Match

Best Local Similarity 39.5%; Score 50; DB 2; Length 953;
Matches 15; Conservative 1; Mismatches 4; Indels 18; Gaps 2;

QY 11 RPPDINPAMWTG-----RGIRPVGRF 31

DB 707 RPPRSNPATYTGAFTHIRDFVELPESKARGYK-VGRF 743

Search completed: April 17, 2001, 15:45:54
Job time: 602 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 17, 2001, 15:48:06 ; Search time 115.07 Seconds
(without alignments)
31.576 Million cell updates/sec

Title: US-09-446-543A-47
Perfect score: 174
Sequence: 1 SRAHQSMETPTDIPNATYGRGIRPVGRF 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database :

SPREMBL.15:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.podent:*
12: sp.unclassified:*
13: sp.vertebrate:*
14: sp.virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	107	61.5	117	13	Q9W624
2	68	39.1	682	2	Q9L8J6
3	61	35.1	1236	2	Q9L8J6
4	58	33.3	1292	2	Q9L8J6
5	56	32.2	428	6	Q9L8J6
6	53	30.5	54	4	Q9L8J6
7	53	30.5	223	2	Q9L8J6
8	53	30.5	465	4	Q9L8J6
9	52	29.9	503	10	Q9L8J6
10	52	29.9	940	2	Q9L8J6
11	51.5	29.6	503	2	Q9L8J6
12	51	29.3	294	5	Q9L8J6
13	51	29.3	428	4	Q9L8J6
14	51	29.3	428	4	Q9L8J6
15	50.5	29.0	176	3	Q9L8J6
16	50.5	29.0	333	2	Q9L8J6
17	50.5	29.0	414	2	Q9L8J6
18	50.5	29.0	538	4	Q9L8J6
19	50.5	29.0	548	10	Q9L8J6

20	50	28.7	863	5	Q9VRY3
21	50	28.7	965	2	Q9PAR9
22	49.5	28.4	306	2	Q9LAD4
23	49	28.2	128	2	Q9LAD4
24	49	28.2	320	2	Q9L7X4
25	49	28.2	419	4	Q9V276
26	49	28.2	772	10	Q9L349
27	49	28.2	1193	2	Q9V5R0
28	49	28.2	1193	2	Q9RPS5
29	49	28.2	1330	2	Q9V284
30	49	28.2	1331	2	Q9V302
31	49	28.2	1379	10	Q9V893
32	49	28.2	1380	10	Q40001
33	49	28.2	1381	10	Q39049
34	49	28.2	1382	10	Q22435
35	49	28.2	1383	10	Q63808
36	48.5	27.9	189	5	Q9VH04
37	48.5	27.9	430	10	Q9SK71
38	48	27.6	157	2	Q9RC1
39	48	27.6	369	4	Q9UG93
40	48	27.6	455	2	Q9V340
41	48	27.6	498	10	Q9STP4
42	48	27.6	517	10	Q9STP3
43	48	27.6	770	4	Q9V284
44	48	27.6	790	10	Q9M371
45	48	27.6	798	4	Q94961

ALIGNMENTS

RESULT 1

Q9W624	PRELIMINARY:	PRT:	117 AA.
ID	Q9W624:		
DR	01-NOV-1999 (TREMBLrel. 12, Created)		
DT	01-NOV-1999 (TREMBLrel. 12, Last sequence update)		
DE	01-NOV-1999 (TREMBLrel. 12, Last annotation update)		
OS	CARASSIUS auratus (Goldfish).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
CC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;		
OX	Cypriniformes; Cyprinidae; Cyprininae; Carassius.		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=BRAIN:		
RA	Saitake H., Minakata H., Fujimoto M.;		
RT	"Carassius Rhamphus" (C-RP amide)."		
RL	Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.		
DR	EMBL: AB020024; BAA76662.1; -		
DE	EMBL: AB020024; BAA76662.1; -		
SO	SEQUENCE 117 AA; 12879 MW; D5DC4CB22038C2B0 CMC64;		

Query Match 61.5%; Score 107; DB 13; Length 117;

Best Local Similarity 57.7%; Pred. No. 4.5e-08;

Matches 15; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 6 HSNETPTDIPNATYGRGIRPVGRF 31

DB 50 HSNETPTDIPNATYGRGIRPVGRF 75

RESULT 2

ID	Q9L8J6	PRELIMINARY:	PRT:	692 AA.
AC	Q9L8J6:			
DT	01-OCT-2000 (TREMBLrel. 15, Created)			
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)			
DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)			
DE	BCHH (FRAGMENT).			
GN	BCHH.			
OS	Rhodospirillum rubrum.			

Query Match 32.2%; Score 56; DB 6; Length 428;
 Best Local Similarity 37.0%; Pred. No. 4.1;
 Matches 10; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 2 RANOSMETERPDPINAWYTGIRPV 28
 DB 195 RANPSLMTYSDCLGATWAGRLKPM 221

RESULT 6
 ID 09UJF9 PRELIMINARY; PRT: 54 AA.

AC 09UJF9; 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE D1479J7.3 (SUSHI-REPEAT PROTEIN (SRPUL)) (FRAGMENT).
 GN D1479J7.3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lawlor S.;
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL035608; CAB55682.1; -.
 FT NON_TER 54
 SO SEQUENCE 54 AA; 6110 MW; E2P3C39F7B961A9F CRC64;

Query Match 30.5%; Score 53; DB 4; Length 54;
 Best Local Similarity 50.0%; Pred. No. 1.2;
 Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 12 TPDINPAMWTGRCIRP 27
 DB 18 TPAYPTWYAGSGYYP 33

RESULT 7
 ID 085605 PRELIMINARY; PRT: 223 AA.

AC 085605; 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE EXCINUCLEASE ABC SUBUNIT A (FRAGMENT).
 GN UVRA.
 OS Prevotella albensis.
 OC Bacteria; CFB group; Bacteroidaceae; Prevotella.
 OX NCBI_TaxID=77768;
 RN [1]
 RP SEQUENCE OF 9-191 FROM N.A.
 RC STRAIN-M384;
 RA Walker N.D., McEwan N.R., Wallace R.J.;
 RT "Prevotella albensis putative uvra gene."
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF070990; AAC24131.2; -.
 DR INTERPRO: IPR001617;
 DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
 KW ATP-binding; Transport.
 FT NON_TER 1
 FT NON_TER 223
 SO SEQUENCE 223 AA; 24579 MW; 2563BA0B0C2996C9 CRC64;

Query Match 30.5%; Score 53; DB 2; Length 223;
 Best Local Similarity 39.5%; Pred. No. 5.5;
 Matches 15; Conservative 1; Mismatches 4; Indels 18; Gaps 2;

QY 11 RTPDINPAMWTG-----RCIRPVGRF 31
 DB 56 RTPRSNPAITYGVFSIDIRLFLVGLPEAKIRGKYP-GRF 92

RESULT 8
 ID 060687 PRELIMINARY; PRT: 465 AA.

AC 060687; 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE SUSHI-REPEAT PROTEIN.
 GN SRPUL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kurosawa H., Inukai T., Inaba T., Goi K., Chang K.-S., Sinjo T.,
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF060367; AAC15765.1; -.
 DR INTERPRO: IPR000436; -.
 DR INTERPRO: IPR001128; -.
 DR PFMW: PF00084; sushi; 3.
 DR PROSITE: PS00086; CYTOCHROME P450; UNKNOWN_1.
 SO SEQUENCE 465 AA; 52971 MW; 4D752B187FF3EFP8 CRC64;

Query Match 30.5%; Score 53; DB 4; Length 465;
 Best Local Similarity 50.0%; Pred. No. 12;
 Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 12 TPDINPAMWTGRCIRP 27
 DB 18 TPAYPTWYAGSGYYP 33

RESULT 9
 ID 09LSC6 PRELIMINARY; PRT: 503 AA.

AC 09LSC6; 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE EMBL/CAB5482.1.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-COLUMBIA;
 RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-COLUMBIA;
 RA Nakamura Y.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
 RT features of the regions of 4,504,864 bp covered by sixty pl and YAC
 RT clones."
 RL DNA Res. 7:131-135 (2000).
 DR EMBL; AB026649; BAB01091.1; -.
 SO SEQUENCE 503 AA; 57928 MW; F5E38CD1BA9C521A CRC64;

Query Match 29.9%; Score 52; DB 10; Length 503;
 Best Local Similarity 33.3%; Pred. No. 18;
 Matches 9; Conservative 8; Mismatches 10; Indels 0; Gaps 0;

OY 3 AHQSHMETRPDINPAMYGRCIRPVG 29
 DB 199 AHQHNLMSFSKDSNSMPCACVPIG 225

RESULT 10
 ID Q9KRW5 PRELIMINARY: PRT: 940 AA.

DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE EXCINUCLEASE ABC, SUBUNIT A.
 GN VC0394.

OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.

NCBI_Taxid=666;
 (1)
 SEQUENCE FROM N.A.

RC STRAIN=EL TOR N16961 / SEROTYPE O1;
 RX MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA McDonald L., Vamathevan J., Bass S., Qin H., Dragol I., Sellers P.,
 RA Salzberg S.L., Utterback T., Fleischmann R.D., Nieman W.C., White O.,
 RA Fraser C.M.,
 RA Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,

RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio

RL Nature 406:477-483(2000).

DR EMBL: AE004127; AAF93567.1; -

DR TIGR: VC0394; -

SO SEQUENCE 940 AA; 104327 MW; 84F93BD9F686F6E2 CRC64;
 Query Match 29.98; Score 52; DB 2; Length 940;
 Best Local Similarity 39.58; Pred. No. 36;
 Matches 15; Conservative 1; Mismatches 4; Indels 18; Gaps 2;

OY 11 RTPDINPAMYG-----RGIKPVGR 31
 DB 694 KFPNSPATYTGITPPIELNAGIOESRSKQYQ-GRP 730

RESULT 11
 ID Q9KRY1 PRELIMINARY: PRT: 503 AA.
 AC Q9KRY1: 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE SUN/NITCEOLAR PROTEIN FAMILY PROTEIN.
 GN VC1502.

OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.

NCBI_Taxid=666;
 (1)

SEQUENCE FROM N.A.
 RC STRAIN=EL TOR N16961 / SEROTYPE O1;
 RX MEDLINE=20406833; PubMed=10952301;

RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA McDonald L., Vamathevan J., Bass S., Qin H., Dragol I., Sellers P.,
 RA Salzberg S.L., Utterback T., Fleischmann R.D., Nieman W.C., White O.,
 RA Fraser C.M.,
 RA Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,

RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio

RL Nature 406:477-483(2000).

DR EMBL: AE004228; AAF94657.1; -

DR TIGR: VC1502; -
 SO SEQUENCE 503 AA; 55997 MW; 2ABD94A2356C9E48 CRC64;

Query Match 29.68; Score 51.5; DB 2; Length 503;
 Best Local Similarity 36.78; Pred. No. 22;
 Matches 11; Conservative 5; Mismatches 13; Indels 1; Gaps 1;

OY 1 SRAQSHMETRPDINPAMYGRCIRPVGR 30
 DB 435 SSSASHSVELDTQAR-EMWGRDVRDEGQ 463

RESULT 12
 ID Q19530 PRELIMINARY: PRT: 294 AA.

DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE F19H6.1 PROTEIN.

GN F19H6.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodermidae; Caenorhabditis.
 NCBI_Taxid=6239;

RP SEQUENCE FROM N.A.

RA McMurray A.;
 RL Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases.

DR EMBL: Z50873; CAA90762.1; -

DR EMBL: Z68115; CAA90762.1; JOINED.

DR EMBL: Z68115; CAA92169.1; -

DR EMBL: Z50873; CAA92169.1; JOINED.

DR HSP: P24941; ICRP.

DR INTERPRO: IPR00719; -

DR INTERPRO: IPR01245; -

DR PIRAM: PF00069; pkinase; 1.

DR PRINTS: PR00109; TYRKINASE.

DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.

DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.

DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.

SO SEQUENCE 294 AA; 34006 MW; 071D4664D049F43 CRC64;

Query Match 29.38; Score 51; DB 5; Length 294;
 Best Local Similarity 44.88; Pred. No. 14;
 Matches 13; Conservative 2; Mismatches 12; Indels 2; Gaps 1;

OY 3 AHQSHMETRPDINPA--WYTGRCIRPVG 29
 DB 139 AHVHSKRIMRDIPKAVFTGIVKLG 167

RESULT 13
 ID Q9U049 PRELIMINARY: PRT: 428 AA.

AC Q9U049: 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DE 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE GANGLIOSIDE SIALIDASE.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

NCBI_Taxid=9606;

RP SEQUENCE FROM N.A.

RC TISSUE=BRAIN;

RA Wada T., Toshikawa Y., Tokuyama S., Kuwabara M., Akita H., Miyagi T.;
 RT "Cloning, expression, and chromosomal mapping of a human ganglioside

RL Biochem. Biophys. Res. Commun. 261:21-27(1999).
 DR EMBL: AB008185; BAA82611.1; -.
 DR INTERPRO: IPR002860; -.
 DR PRAM: PF02012; BMR: 3.
 SO SEQUENCE 428 AA; 48252 MW; 35D1DD9359A78C98 CRC64;

Query Match 29.3%; Score 51; DB 4; Length 428;
 Best Local Similarity 33.3%; Pred. No. 21;

Matches 9; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

OY 2 RAHQSMETRTPDINPAMTGRGIRPV 28

DB 195 KTRPHSLMITYSDDLGVTHGRLIRPM 221

RESULT 14

O9NOE1 PRELIMINARY; PRT: 629 AA.

AC O9NOE1

DT 01-OCT-2000 (TREMBlrel. 15, Created)

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)

DE NURAMINIDASE (EC 3.2.1.18).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;

RN 11

RP SEQUENCE FROM N.A.

RA Monti E., Bassi M.T., Papini N., Riboldi M., Manzoni M., Venerando B.,

RA Croci G., Preti A., Balabio A., Tettamanzi G., Borsani G.;

RT Identification and expression of NEU3, a novel human stialidase

RT associated to the plasma membrane.";

RL Biochem. J. 349:343-351(2000).

KW EMBL: Y18563; CAB96131.1; -.

KW Hydrolase; Glycosidase.

SO SEQUENCE 629 AA; 69702 MW; 97C5464B70E69B4B CRC64;

Query Match 29.3%; Score 51; DB 4; Length 629;

Best Local Similarity 33.3%; Pred. No. 32;

Matches 9; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

OY 2 RAHQSMETRTPDINPAMTGRGIRPV 28

DB 396 KTRPHSLMITYSDDLGVTHGRLIRPM 422

RESULT 15

O08689 PRELIMINARY; PRT: 176 AA.

AC O08689

DT 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DE CHROMOSOME XV READING FRAME ORF YOR253W.

OS Saccharomyces cerevisiae (Baker's Yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;

OC Saccharomycetaceae; Saccharomyces.

OX NCBI_TaxID=4932;

RN 11

RP SEQUENCE FROM N.A.

RA MIP5;

RL Submitted (JUL-1996) to the EMBL/Genbank/DBJ databases.

RN 12

RP SEQUENCE FROM N.A.

RX MEDLINE=97298311; PubMed=9153759;

RA Jauniaux J.C., Polirey R.;

RT Sequencing analysis of a 36.8 kb fragment of yeast chromosome XV

RT reveals 26 open reading frames including SEC63, COC31, SUG2, GCD1,

RT RDL2, PWT1, PRC1 and VPH1.";

RL Yeast 13:483-487(1997).

DR EMBL: Z75161; CAA99475.1; -.
 DR INTERPRO: IPR000182; -.
 DR INTERPRO: IPR000345; -.
 DR PRAM: PF00583; Acetyltransf. 1.
 DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1
 SO SEQUENCE 176 AA; 19727 MW; 4F09C597A690BA0 CRC64;

Query Match 29.0%; Score 50.5; DB 3; Length 176;

Best Local Similarity 31.2%; Pred. No. 9.8;

Matches 10; Conservative 4; Mismatches 15; Indels 3; Gaps 1;

OY 1 SRAHQSMETRTPDINPAMTGRGIRPV 29

DB 120 SECHQHNVFYLLPAVDDITKQMTAAGPEQYG 151

Search completed: April 17, 2001, 15:48:07
 Job time: 560 sec

Tue Apr 17 15:46:12 2001

us-09-446-543a-47.rpt

Page 6

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 17, 2001, 15:48:47 : Search time 39.1 Seconds
(without alignments)
27.159 Million cell updates/sec

Title: US-09-446-543A-47

Sequence: 1 SRAHSHMETRPDPINPAMWTGIRPVGRF 31

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 93435 segs, 3425486 residues
Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	174	100.0	83	1 PRP_PAT	P81278 Rattus norv
2	163	93.7	98	1 PRP_HOVIN	P81264 Bos taurus
3	149	85.6	87	1 PRP_HUMAN	P81277 homo sapien
4	56	32.2	962	1 UVRA_METH	O26543 methanobact
5	53	30.5	798	1 UNR_PAT	P18395 rattus norv
6	52	29.9	972	1 UVRA_MYCTU	P94972 mycobacteri
7	52	29.9	873	1 UVRA_RHINE	P56899 rhizobium m
8	51	29.3	943	1 UVRA_HAEIN	P44410 haemophilus
9	50	28.7	264	1 Y355_BUCAI	P57436 buchnera ap
10	50	28.7	482	1 R167_YEAST	P39743 saccharomyc
11	50	28.7	940	1 UVRA_ECOLI	P07671 escherichia
12	50	28.7	941	1 UVRA_SALTY	P37434 salmonella
13	50	28.7	953	1 UVRA_RICPR	O92633 rickettsia
14	49.5	28.4	374	1 YHHO_ECOLI	P31993 escherichia
15	49	28.2	952	1 UVRA_THETH	O56242 thetus aqu
16	49	28.2	970	1 UVRA_SYNX3	P73412 streptocyst
17	49	28.2	1014	1 UVRA_SPRCO	O92507 streptomyces
18	48	27.6	569	1 UVRA_VITST	O08518 vitreoscilli
19	48	27.6	719	1 ARP_YEAST	P32770 saccharomyc
20	48	27.6	798	1 UNR_HUMAN	O75534 homo sapien
21	48	27.6	950	1 UVRA_NEIGO	O50968 nelisseria g
22	48	27.6	960	1 UVRA_TREPA	O83527 treponema p
23	48	27.6	1194	1 BCNH_RHOCA	P26162 rhodobacter
24	47.5	27.3	501	1 TRA2_MOUSE	P39429 mus musculu
25	47.5	27.3	941	1 GCSP_MYCTU	O50601 mycobacteri
26	47	27.0	453	1 TBB2_GEOCN	P32925 geotrichum
27	47	27.0	926	1 UVRA_AQDAE	O66311 aquifex aeo
28	47	27.0	1083	1 T2D3_HUMAN	O02668 homo sapien
29	46.5	26.7	272	1 TRA2_DROVI	O02008 drosophila
30	46.5	26.7	652	1 TETP_CLOPE	O46306 clostridium
31	46	26.4	332	1 GRP2_MOUSE	O89100 m grb2-rela
32	46	26.4	330	1 GRP2_HUMAN	O75791 h grb2-rela
33	46	26.4	925	1 UVRA_ZYMOO	O31151 zymomonas m

34	46	26.4	957	1 UVRA_BACSU	O34863 bacillus su
35	46	26.4	992	1 UVRA_MICLU	P13567 micrococcus
36	46	26.4	1016	1 UVRA_DETRA	O46577 delnoccocus
37	45.5	26.1	486	1 DMPC_PSESP	P19059 pseudomonas
38	45	25.9	126	1 YBSG_ECOLI	P75758 escherichia
39	45	25.9	184	1 RL6_THEMA	O92ae4 thermotoga
40	45	25.9	430	1 YAE9_SCHPO	O09849 schizosacch
41	45	25.9	676	1 EXL1_HUMAN	O92935 homo sapien
42	45	25.9	773	1 DROL_THRGO	P56689 thermococcu
43	45	25.9	817	1 DRPI_YEAST	P37370 saccharomyc
44	45	25.9	1523	1 DPOL_THEFM	P74918 thermococcu
45	44.5	25.6	240	1 PLSC_HELPY	O23903 helicobacte

ALIGNMENTS

RESULT	1	STANDARD	PRT	83 AA.
PRP_PAT	PRP_PAT			
ID	P81278			
AC	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DE	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	PROLACTIN-RELEASING PEPTIDE PRECURSOR (PRP) (PROLACTIN-RELEASING			
DE	HORMONE) [CONTAINS: PROLACTIN-RELEASING PEPTIDE PRP31; PROLACTIN-			
DE	RELEASING PEPTIDE PRP20].			
GN	PRP.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_Taxid=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RA	MEDLINE=96266781; Pubmed=9607765;			
RA	Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,			
RA	Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,			
RA	Kurokawa T., Nishimura O., Onda H., Fujino M.;			
RT	"A prolactin-releasing peptide in the brain.";			
RL	Nature 393:272-276(1998).			
CC	FUNCTION: STIMULATES PROLACTIN (PR) RELEASE AND REGULATES THE			
CC	EXPRESSION OF PROLACTIN THROUGH ITS RECEPTOR HGR3. MAY STIMULATE			
CC	LACTOROPHUS DIRECTLY TO SECRETE PRL.			
CC	- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.			
CC	- This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	EMBL; AB015418; BAA29026.1;			
DR	Hormone; Amidation; Signal; Cleavage on pair of basic residues.			
KW	SIGNAL			
FT	PEPTIDE 22 21			
FT	PEPTIDE 33 52			
FT	PEPTIDE 52 52			
FT	MOD_RES 33 52			
SO	SEQUENCE 83 AA; 9215 MW; DQC75A264EEB4F29 CRC64;			

Query Match 100.0%; Score 174; DB 1; Length 83;
Best Local Similarity 100.0%; Pred. No. 2e-18;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRAHSHMETRPDPINPAMWTGIRPVGRF 31
DB 22 SRAHSHMETRPDPINPAMWTGIRPVGRF 52

RESULT 2


```

CC (ABC TRANSPORTERS). CONTAINS TWO ABC DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE000828; AAB84949.1; -
CC InterPro: IPR001617; -
CC Pfam: PF00005; ABC_tran. 2.
CC PROSITE: PS00211; ABC_TRANSPORTER; 2
CC SOS response; Excision nuclease; DNA repair; ATP-binding; Repeat;
CC DNA-binding; Zinc-finger.
CC NP_BIND 38 45 ATP (POTENTIAL).
CC NP_BIND 649 656 ATP (POTENTIAL).
CC ZN_FING 748 774 C4-TYPE.
CC SEQUENCE 962 AA; 108395 MW; 2C0EF7FC41CDD060 CRC64;

Query Match
Best Local Similarity 42.1%; Score 56; DB 1; Length 962;
Matches 16; Conservative 0; Mismatches 4; Indels 18; Gaps 2;

QY 11 RPPDIPAWYTG-----RGTRPGGRF 31
DB 703 RPRSNPARYTGVEFHIRELFQTPERRKRGYP-GRF 739

RESULT 5
UNR_RAT STANDARD; PRT; 798 AA.
ID UNR_RAT
AC P18395;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE UNR PROTEIN.
GN UNR.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA MEDLINE-90370473; PubMed-2204029;
RA Jeffers M., Pacinucci R., Pellicer A.;
RA "Characterization of unr: a gene closely linked to N-ras.";
RA Nucleic Acids Res. 18:4891-4899(1990).
CC 1- FUNCTION: RNA-BINDING PROTEIN (BY SIMILARITY).
CC 1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC 1- SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X52311; CAA36549.1; -
CC PIR: S11210; S11210.
CC HSP: P15277; IMC.
CC InterPro: IPR002059; -
CC Pfam: PF00313; CSD; 8.
CC PROSITE: PS00352; COLD_SHOCK; 4.
CC RNA-binding; Repeat.
CC NP_BIND 26 87 CSD 1.
CC NP_BIND 136 179 CSD 2 (INCOMPLETE).
CC DOMAIN 186 245 CSD 3.
CC FT
CC DOMAIN

```

```

FT DOMAIN 297 337 CSD 4 (INCOMPLETE).
FT DOMAIN 349 410 CSD 5.
FT DOMAIN 447 507 CSD 6.
FT DOMAIN 519 579 CSD 7.
FT DOMAIN 610 670 CSD 8.
FT DOMAIN 674 735 CSD 9.
SQ SEQUENCE 798 AA; 88894 MW; F484B3FA8B095A4 CRC64;

Query Match
Best Local Similarity 43.5%; Score 53; DB 1; Length 798;
Matches 10; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 6 HSMETRPDIPAWYTGIRPV 28
DB 583 HSYNGITEANPTISGKVRPL 605

RESULT 6
UNR_MYCTU STANDARD; PRT; 972 AA.
ID UNR_MYCTU
AC P94972;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE EXCINUCLEASE ABC SUBUNIT A.
GN UNR OR RV1638 OR MYC06H11.02.
OS Mycobacterium tuberculosis.
CC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
CC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_Taxid=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RA MEDLINE-98295987; PubMed-9634230;
RA Core S.T., Brosch R., Parish J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock R., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holtroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulston J.E.,
RA Taylor K., Whitehead S., Barrell B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RA complete genome sequence.";
RA Nature 393:537-544(1998).
CC 1- FUNCTION: THE ABC EXCISION NUCLEASE IS A DNA REPAIR ENZYME THAT
CC CATALYZES THE EXCISION REACTION OF UV-DAMAGED NUCLEOTIDE SEGMENTS
CC PRODUCING OLIGOMERS HAVING THE MODIFIED BASE(S). UNR IS AN ATPASE
CC AND A DNA-BINDING PROTEIN THAT PREFERENTIALLY BINDS SINGLE-
CC STRANDED OR UV-IRRADIATED DOUBLE-STRANDED DNA (BY SIMILARITY).
CC 1- SUBUNIT: CONSISTS OF THREE SUBUNITS: UNR, UNRB AND UNVC.
CC 1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC 1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC (ABC TRANSPORTERS). CONTAINS TWO ABC DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z85982; CAB06633.1; -
CC TubercuList; Rv1638; -
CC InterPro: IPR001617; -
CC Pfam: PF00005; ABC_tran. 2.
CC PROSITE: PS00211; ABC_TRANSPORTER; 2.
CC SOS response; Excision nuclease; DNA repair; ATP-binding; Repeat;
CC DNA-binding; Zinc-finger.
CC NP_BIND 32 39 ATP (POTENTIAL).
CC NP_BIND 654 661 ATP (POTENTIAL).
CC FT
CC NP_BIND

```

FT ZN_FING 257 285 C4-TYPE (ATYPICAL).
 FT ZN_FING 753 779 C4-TYPE.
 SO SEQUENCE 972 AA; 106131 MM; 8937A764E592D901 CRC64;

Query Match 29.9%; Score 52; DB 1; Length 972;
 Best Local Similarity 39.5%; Pred. No. 9.4;
 Matches 15; Conservative 1; Mismatches 4; Indels 18; Gaps 2;

OY 11 RTDINPAMYTG-----RGIRPYGRF 31
 DB 708 RTPRSNPATYTGVDKIRPLFAATEAKYRGYGP-GRF 744

RESULT 7
 ID UVRA_RHIME STANDARD; PRT: 973 AA.
 AC P56899;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE EXCINUCLEASE ABC SUBUNIT A.
 GN UVRA.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RA Gailbert F., Capela D., Hubler-Barloy F., Gattus M., Batut J.,
 RA Bolstad P., Gouzy J., Kahn D., Thebaud P., Coffeau A.,
 RA Purnelle B., Pohl T., Botte G., Schneider S., Portetelle D.,
 RA Vandewol M., Puhler A., Becker A., Weidner S.,
 RL Submitted (MAY-2000) to the SWISS-PROT data bank.
 RN 12
 RP SEQUENCE OF 1-140 FROM N.A.
 RC STRAIN=2021;
 RX MEDLINE=99430868; PubMed=10503543;
 RA Taplas A., Barbe J.;
 RT "Regulation of divergent transcription from the uvra-ssb promoters in
 RT Sinorhizobium meliloti";
 RL Mol. Gen. Genet. 262:121-130(1999).
 CC -1- FUNCTION: THE ABC EXCISION NUCLEASE IS A DNA REPAIR ENZYME THAT
 CC CATALYZES THE EXCISION REACTION OF UV-DAMAGED NUCLEOTIDE SEGMENTS
 CC PRODUCING OLIGOMERS HAVING THE MODIFIED BASE(S). UVRA IS AN ATPASE
 CC AND A DNA-BINDING PROTEIN THAT PREFERENTIALLY BINDS SINGLE-
 CC STRANDED OR UV-IRRADIATED DOUBLE-STRANDED DNA (BY SIMILARITY).
 CC -1- SUBUNIT: CONSISTS OF THREE SUBUNITS: UVRA, UVRA AND UVRC.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
 CC (ABC TRANSPORTERS). CONTAINS TWO ABC DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@lsb-sib.ch).
 CC
 CC EMBL: AF125162; AF03210.1;
 DR Interpro: IPR001617;
 DR PROSITE: PS00211; ABC_TRANSPORTER; PARTIAL.
 KM SOS response; Excision nuclease; DNA repair; ATP-binding; Repeat;
 KM DNA-binding; Zinc-finger.
 FT NP_BIND 34 41 ATP (POTENTIAL).
 FT NP_BIND 662 669 ATP (POTENTIAL).
 FT ZN_FING 761 787 C4-TYPE.
 FT ZN_FING 19 19 G -> A (IN REF. 2).
 FT CONFLICT 67 67 F -> S (IN REF. 2).
 FT CONFLICT 67 67 F -> S (IN REF. 2).
 SO SEQUENCE 973 AA; 107191 MM; 3E1A8B14527A47FE CRC64;

Query Match 29.9%; Score 52; DB 1; Length 973;
 Best Local Similarity 39.5%; Pred. No. 9.4;
 Matches 15; Conservative 1; Mismatches 4; Indels 18; Gaps 2;

OY 11 RTDINPAMYTG-----RGIRPYGRF 31
 DB 716 RTPRSNPATYTGCAFTPIRDFAGLPEAKAKYGP-GRF 752

RESULT 8
 ID UVRA_HAEIN STANDARD; PRT: 943 AA.
 AC P44410; Q48151;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE EXCINUCLEASE ABC SUBUNIT A.
 GN UVRA OR H10249.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Haemophilus.
 OX NCBI_TaxID=727;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=RD / KW20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7342800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kervage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Utermack T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhmann J.L., Geoghagen N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small R.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus Influenzae
 RT Rd";
 RL Science 269:496-512(1995).
 RN 12
 RP SEQUENCE FROM N.A.
 RC STRAIN=NTHI N187;
 RX MEDLINE=97080495; PubMed=8921840;
 RA de la Morena M.L., Hendrixson D.R., St Geme J.W., III;
 RT "Isolation and characterization of the Haemophilus Influenzae uvra
 RT gene";
 RL Gene 177:23-28(1996).
 RN 13
 RP SEQUENCE OF 1-71 FROM N.A.
 RC STRAIN=NTHI TN106;
 RX MEDLINE=94341556; PubMed=8063092;
 RA Jarosik G.P., Hansen E.J.;
 RT "Cloning and sequencing of the Haemophilus influenzae ssb gene
 RT encoding single-strand DNA-binding protein";
 RL Gene 146:101-103(1994).
 CC -1- FUNCTION: THE ABC EXCISION NUCLEASE IS A DNA REPAIR ENZYME THAT
 CC CATALYZES THE EXCISION REACTION OF UV-DAMAGED NUCLEOTIDE SEGMENTS
 CC PRODUCING OLIGOMERS HAVING THE MODIFIED BASE(S). UVRA IS AN ATPASE
 CC AND A DNA-BINDING PROTEIN THAT PREFERENTIALLY BINDS SINGLE-
 CC STRANDED OR UV-IRRADIATED DOUBLE-STRANDED DNA (BY SIMILARITY).
 CC -1- SUBUNIT: CONSISTS OF THREE SUBUNITS: UVRA, UVRA AND UVRC.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
 CC (ABC TRANSPORTERS). CONTAINS TWO ABC DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@lsb-sib.ch).
 CC

FT DOMAIN 421 482 SH3.
 SO SEQUENCE 482 AA; 52774 MW; 3FOAB53EBC95A5B CRC64;
 Query Match 28.7%; Score 50; DB 1; Length 482;
 Best Local Similarity 50.0%; Pred. No. 8.8;
 Matches 11; Conservative 3; Mismatches 2; Indels 6; Gaps 2;
 0Y 11 RTPDINPARYTGR-----GIRP 27
 DB 453 RTPDVN-EMWTRNGQGVFP 473
 RESULT 11
 UVR A ECOLI
 ID UVR A ECOLI STANDARD; PRT; 940 AA.
 AC P07671; P76788;
 DT 01-APR-1988 (rel. 07, Created)
 DT 01-APR-1988 (rel. 07, Last sequence update)
 DT 30-MAY-2000 (rel. 39, Last annotation update)
 DE EXCINUCLEASE ABC SUBUNIT A.
 GN UVR A OR DINE.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=86168204; Pubmed=3007478;
 RA Husein I., van Houten B., Thomas D.C., Sancar A.;
 RT "Sequences of Escherichia coli uvrA gene and protein reveal two
 RT potential ATP binding sites.";
 RL J. Biol. Chem. 261:4895-4901(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA STRAIN-K12 / NG1655;
 RA MEDLINE=94089392; Pubmed=8265357;
 RA Blatner F.R., Burland V.D., Plunkett G. III, Sofia H.J.,
 RA Daniels D.L.;
 RT "Analysis of the Escherichia coli genome. IV. DNA sequence of the
 RT region from 89.2 to 92.8 minutes.";
 RL Nucleic Acids Res. 21:5408-5417(1993).
 RN [3]
 RP SEQUENCE OF I-25 FROM N.A.
 RA MEDLINE=83299251; Pubmed=6310514;
 RA Beckendorf C., Brandma J.A., Kartasova T., van de Putte P.;
 RT "In vivo regulation of the uvrA gene: role of the '-10' and '-35'
 RT promoter regions.";
 RL Nucleic Acids Res. 11:5795-5810(1983).
 RN [4]
 RP SEQUENCE OF I-14 FROM N.A.
 RA MEDLINE=82220077; Pubmed=6283374;
 RA Sancar A., Sancar G.B., Rupp W.D., Little J.W., Mount D.W.;
 RT "LexA protein inhibits transcription of the E. coli uvrA gene in
 RT vitro.";
 RL Nature 298:96-98(1982).
 RN [5]
 RP CHARACTERIZATION.
 RA MEDLINE=91208117; Pubmed=1826851;
 RA Myles G.M., Sancar A.;
 RT "Isolation and characterization of functional domains of UVR A.";
 RL Biochemistry 30:3834-3840(1991).
 RN [6]
 RP MUTAGENESIS OF CYS-253.
 RA MEDLINE=89380205; Pubmed=2550431;
 RA Navaratnam S., Myles G.N., Strange R.W., Sancar A.;
 RT "Evidence from extended X-ray absorption fine structure and site-
 RT specific mutagenesis for zinc fingers in UVR A protein of Escherichia
 RT coli.";
 RL J. Biol. Chem. 264:16067-16071(1989).
 CC -1- FUNCTION: THE ABC EXCISION NUCLEASE IS A DNA REPAIR ENZYME THAT
 CC CATALYZES THE EXCISION REACTION OF UV-DAMAGED NUCLEOTIDE SEGMENTS
 CC PRODUCING OLIGOMERS NAVING THE MODIFIED BASE(S). UVR A IS AN ATPASE

CC AND A DNA-BINDING PROTEIN THAT PREFERENTIALLY BINDS SINGLE-
 CC STRANDED OR UV-IRRADIATED DOUBLE-STRANDED DNA.
 CC -1- SUBUNIT: CONSISTS OF THREE SUBUNITS: UVR A, UVR B AND UVR C.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
 CC -1- MISCELLANEOUS: BINDS ABOUT 2 ZINC ATOMS/MOLECULE.
 CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
 CC (ABC TRANSPORTERS). CONTAINS TWO ABC DOMAINS.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M13495; AAA24754.1; -;
 DR EMBL; U00006; AAC43152.1; -;
 DR EMBL; AE000479; AAC77028.1; -;
 DR EMBL; X01621; CAA25764.1; -;
 DR EMBL; J01721; AAA24753.1; -;
 DR PIR; A23869; BVECOA.
 DR ECO2DBASE; H124.0; 6TH EDITION.
 DR EcoGene; EG11061; UVR A.
 DR InterPro; IPR001617; -;
 DR Pfam; PF00005; ABC_tran; 2.
 DR PROSITE; PS00211; ABC_TRANSPORTER; 2.
 KW SOS response; Excision nuclease; DNA repair; ATP-binding; Repeat;
 FT DNA-binding; Zinc-finger.
 FT NP_BIND 31 38 ATP.
 FT NP_BIND 640 647 ATP.
 FT ZN_FING 253 280 C4-TYPE.
 FT ZN_FING 740 766 C4-TYPE.
 FT MOTACEN 253 253 C->A,H,S: REDUCED ACTIVITY.
 SQ SEQUENCE 940 AA; 103867 MW; D61AAEB514B860C CRC64;
 Query Match 28.7%; Score 50; DB 1; Length 940;
 Best Local Similarity 39.5%; Pred. No. 18;
 Matches 15; Conservative 0; Mismatches 5; Indels 18; Gaps 2;
 0Y 11 RTPDINPARYTGR-----GIRPVGRE 31
 DB 695 RTPRSNPARYTGVTFVRELPAVGPSRARGYRP-GRF 731
 RESULT 12
 UVR A SALTY
 ID UVR A SALTY STANDARD; PRT; 941 AA.
 AC P37434;
 DT 01-OCT-1994 (rel. 30, Created)
 DT 01-OCT-1994 (rel. 30, Last sequence update)
 DT 30-MAY-2000 (rel. 39, Last annotation update)
 DE EXCINUCLEASE ABC SUBUNIT A.
 GN UVR A.
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Salmonella.
 OX NCBI_TaxID=602;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Alberti M., Li Y.F., Sancar A., Hearst J.E.;
 RT Submitted (Aug-1992) to the EMBL/Genbank/DBJ databases.
 RL -1- FUNCTION: THE ABC EXCISION NUCLEASE IS A DNA REPAIR ENZYME THAT
 CC CATALYZES THE EXCISION REACTION OF UV-DAMAGED NUCLEOTIDE SEGMENTS
 CC PRODUCING OLIGOMERS HAVING THE MODIFIED BASE(S). UVR A IS AN ATPASE
 CC AND A DNA-BINDING PROTEIN THAT PREFERENTIALLY BINDS SINGLE-
 CC STRANDED OR UV-IRRADIATED DOUBLE-STRANDED DNA (BY SIMILARITY).
 CC -1- SUBUNIT: CONSISTS OF THREE SUBUNITS: UVR A, UVR B AND UVR C.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
 CC -1- MISCELLANEOUS: BINDS ABOUT 2 ZINC ATOMS/MOLECULE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
 CC (ABC TRANSPORTERS). CONTAINS TWO ABC DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: M93014; AAA27250.1; -
DR StyGene: SG10413; uvra.
DR InterPro: IPR001617; -
DR Pfam: PF00005; ABC_tran; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
KW SOS response; Excision nuclease; DNA repair; ATP-binding; Repeat;
KW DNA-binding; zinc-finger.
FT NP_BIND 31 38 ATP.
FT ZN_BIND 640 647 ATP.
FT ZN_FING 253 280 C4-TYPE.
FT ZN_FING 740 766 C4-TYPE.
SQ SEQUENCE 941 AA; 103928 MW; CAAFC9F549060C26 CRC64;

Query Match 28.7%; Score 50; DB 1; Length 941;
Best Local Similarity 39.5%; Pred. No. 18;
Matches 15; Conservative 0; Mismatches 5; Indels 18; Gaps 2;

QY 11 RTPDINPAMVYG-----RGTRPGVGRF 31
DB 695 RTPRSNPATYTGATFTRDVEYELPESKRGYK-GRF 731

RESULT 13
ID UVRA_RICPR STANDARD; PRT; 953 AA.
AC 092CC3;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE EXCISELEASE ABC SUBUNIT A.
GN UVRA OR RP835.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=782;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MADRID E;
RX MEDLINE=99039499; PubMed=9823893;
RA Andersson S.G.E., Zomrodipour A., Andersson J.O.,
RA Sticheritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of
RT mitochondria.";
RT Nature 396:133-140(1998).
CC -1- FUNCTION: THE ABC EXCISION NUCLEASE IS A DNA REPAIR ENZYME THAT
CC CATALYZES THE EXCISION REACTION OF UV-DAMAGED NUCLEOTIDE SEGMENTS
CC PRODUCING OLIGOMERS HAVING THE MODIFIED BASE(S). UVRA IS AN ATPASE
CC AND A DNA-BINDING PROTEIN THAT PREFERENTIALLY BINDS SINGLE-
CC STRANDED OR UV-IRRADIATED DOUBLE-STRANDED DNA (BY SIMILARITY).
CC -1- SUBUNIT: CONSISTS OF THREE SUBUNITS: UVRA, UVRA AND UVRC.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE ABC-BINDING TRANSPORT PROTEIN FAMILY
CC (ABC TRANSPORTERS). CONTAINS TWO ABC DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: A1235273; CAA15260.1; -
DR InterPro: IPR001617; -
DR Pfam: PF00005; ABC_tran; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
KW SOS response; Excision nuclease; DNA repair; ATP-binding; Repeat;
KW DNA-binding; zinc-finger.
FT NP_BIND 33 40 ATP (POTENTIAL).
FT NP_BIND 652 659 ATP (POTENTIAL).
FT ZN_FING 752 778 C4-TYPE.
SQ SEQUENCE 953 AA; 106287 MW; 6209A66241379421 CRC64;

Query Match 28.7%; Score 50; DB 1; Length 953;
Best Local Similarity 39.5%; Pred. No. 18;
Matches 15; Conservative 1; Mismatches 4; Indels 18; Gaps 2;

QY 11 RTPDINPAMVYG-----RGTRPGVGRF 31
DB 707 RTPRSNPATYTGATFTRDVEYELPESKRGYK-GRF 743

RESULT 14
ID YHND_ECOLI STANDARD; PRT; 374 AA.
AC P31993; P76703;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOTHETICAL 41.1 KDA PROTEIN IN RHB-PIT INTERGENIC REGION.
GN YHNT.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=94316500; PubMed=8041620;
RA Sofia H.J., Burland V., Daniels D.L., Plunkett G. III, Blattner F.R.;
RT "Analysis of the Escherichia coli genome. V. DNA sequence of the
RT region from 76.0 to 81.5 minutes.";
RT Nucleic Acids Res. 22:2576-2586(1994).
RN [2]
RP SEQUENCE OF 225-374 FROM N.A.
RC STRAIN=K12;
RX MEDLINE=93259920; PubMed=8387990;
RA Zhao S., Sandt C.H., Feulner G., Vlazny D.A., Gray J.A., Hill C.W.;
RT "Rns elements of Escherichia coli K-12: complex composites of shared
RT and unique components that have different evolutionary histories.";
RL J. Bacteriol. 175:2799-2808(1993).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
CC (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE ABC-2 SUBFAMILY OF INTEGRAL MEMBRANE
CC PROTEINS.
CC -1- SIMILARITY: TO E-COLI YBHR AND YBHS AND TO THE C-TERMINAL OF
CC E-COLI YHIC.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: U00039; AAB18460.1; ALT_INIT.
DR EMBL: AE000424; AAC76510.1; ALT_INIT.
DR EMBL: L02370; AAC61886.1; -
DR EcoGene: EG11767; yhnD.
DR InterPro: IPR000412; -
DR PROSITE: PS00890; ABC2_MEMBRANE; 1.
KW Hypothetical protein; Transport; Transmembrane; Inner membrane.
FT TRANSMEM 23 43 POTENTIAL.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OH protein - protein search, using sw model

Run on: April 17, 2001, 15:39:47 ; Search time 61.54 Seconds
(without alignments)
9.677 Million cell updates/sec

Title: US-09-446-543a-47
Perfect score: 174
Sequence: 1 SRAHSHMETRPDINPAMYTGKIRPVGFR 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*

1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*

2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*

3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*

4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*

5: /cgn2_6/ptodata/2/1aa/PCUS.COMB.pep:*

6: /cgn2_6/ptodata/2/1aa/backfilltest.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	174	100.0	31	US-09-105-678A-8	Sequence 8, App1
2	174	100.0	31	US-09-105-678A-37	Sequence 37, App1
3	174	100.0	31	US-09-172-353-4	Sequence 4, App1
4	174	100.0	32	US-09-105-678A-38	Sequence 38, App1
5	174	100.0	33	US-09-105-678A-39	Sequence 39, App1
6	163	93.7	31	US-09-105-678A-7	Sequence 7, App1
7	163	93.7	31	US-09-105-678A-31	Sequence 31, App1
8	163	93.7	32	US-09-105-678A-32	Sequence 32, App1
9	163	93.7	33	US-09-105-678A-33	Sequence 33, App1
10	152	87.4	29	US-09-105-678A-29	Sequence 29, App1
11	149	85.6	31	US-09-105-678A-9	Sequence 9, App1
12	149	85.6	31	US-09-105-678A-43	Sequence 43, App1
13	149	85.6	32	US-09-105-678A-44	Sequence 44, App1
14	149	85.6	33	US-09-105-678A-45	Sequence 45, App1
15	146	66.7	20	US-09-105-678A-40	Sequence 40, App1
16	146	66.7	21	US-09-105-678A-41	Sequence 41, App1
17	146	66.7	22	US-09-105-678A-42	Sequence 42, App1
18	146	66.7	20	US-09-105-678A-34	Sequence 34, App1
19	146	66.7	21	US-09-105-678A-35	Sequence 35, App1
20	146	66.7	22	US-09-105-678A-36	Sequence 36, App1
21	105	60.3	19	US-09-105-678A-30	Sequence 30, App1
22	105	60.3	20	US-09-105-678A-46	Sequence 46, App1
23	105	60.3	21	US-09-105-678A-47	Sequence 47, App1
24	105	60.3	22	US-09-105-678A-48	Sequence 48, App1
25	104	58.8	21	US-09-105-678A-28	Sequence 28, App1
26	51	29.3	239	US-08-712-709-7	Sequence 7, App1
27	51	29.3	239	US-09-049-671-3	Sequence 3, App1

28	51	29.3	239	3	US-09-295-068-3	Sequence 3, App1
29	51	29.3	239	3	US-09-111-444-7	Sequence 7, App1
30	50	28.7	349	1	US-08-118-270-71	Sequence 71, App1
31	50	28.7	349	5	PCR-US93-08528-71	Sequence 71, App1
32	50	28.7	940	4	US-09-078-347A-1	Sequence 1, App1
33	49	28.2	774	3	US-08-902-632-2	Sequence 2, App1
34	49	28.2	774	3	US-09-073-354-1	Sequence 1, App1
35	49	28.2	774	3	US-08-656-005A-1	Sequence 1, App1
36	49	28.2	774	4	US-09-073-259-1	Sequence 1, App1
37	49	28.2	774	4	US-09-363-095-1	Sequence 1, App1
38	47.5	27.3	80	2	US-08-691-814B-31	Sequence 31, App1
39	47.5	27.3	501	1	US-08-331-394-4	Sequence 4, App1
40	47.5	27.3	501	1	US-08-250-858-4	Sequence 4, App1
41	47.5	27.3	501	1	US-08-446-915-4	Sequence 4, App1
42	47.5	27.3	501	2	US-08-744-139-4	Sequence 4, App1
43	47.5	27.3	501	5	PCR-US95-06639-4	Sequence 4, App1
44	46	26.4	330	2	US-08-815-176-1	Sequence 1, App1
45	46	26.4	335	2	US-08-982-232-7	Sequence 7, App1

ALIGNMENTS

RESULT 1
US-09-105-678A-8
Sequence 8, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105, 678A
FILING DATE: 26-JUN-1998
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27, 026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-8

Query Match 100.0%; Score 174; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 5.5e-20;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 SRAHQSMETRTPDINPAMYTGKGRPVGRF 31
      |||||
DB 1 SRAHQSMETRTPDINPAMYTGKGRPVGRF 31

```

RESULT 2
US-09-105-678A-37

Query Match	100.0%;	Score 174;	DB 3;	Length 31;
Best Local Similarity	100.0%;	Pred. No. 5.5e-20;		
Matches	31;	Conservative	0;	Mismatches 0;
				Indels 0;
				Gaps 0;

```

RESULT      3
US-09-172-353-4
; Sequence 4, Application US/09172353
; Patent No. 6197530
; GENERAL INFORMATION:
; APPLICANT: Stricker-Kongra, Alain
; APPLICANT: Gu, Wei
; TITLE OF INVENTION: GPR10 AS A TARGET FOR IDENTIFYING WEIGHT MODULATING COMPOUNDS
; FILE REFERENCE: 0734/102001
; CURRENT APPLICATION NUMBER: US/09/172,353
; CURRENT FILING DATE: 1998-10-14
; NUMBER OF SEQ. ID NOS.: 7
; SOFTWARE: FastSeq for Windows Version 3.0

```

```

; SEQ ID NO 4
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-172-353-4

```

Query Match	100.0%	Score 174	DB 4	Length 31
Best Local Similarity	100.0%	Pred. No. 5	5e-20	
Matches 31	Conservative 0	Mismatches 0	Indels 0	Gaps 0

RESULT 4
 US-09-105-678A-38
 : Sequence 38, Application US/09105678A
 : Patent No. 6103862
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Suenaga, Masato
 : APPLICANT: Moriya, Takeo
 : APPLICANT: Tanaka, Yoko
 : APPLICANT: Nishimura, Osamu
 :
 : TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
 :
 : NUMBER OF SEQUENCES: 52
 :
 : CORRESPONDENCE ADDRESS:
 : ADDRESS: DILE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
 : STREET: 130 Water Street
 : CITY: Boston
 : STATE: MA
 : COUNTRY: USA

Query Match	100.0%	Score 174:	DB 3:	Length 32:
Best local Similarity	100.0%	Pred. No. 5.7e-20:		
Matches	31;	Conservative	0;	Indels 0; Gaps 0;
		Mismatches	0;	

RESULT 5
US-09-105-678A-39

Sequence 39, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-39

Query Match 100.0%; Score 174; DB 3; Length 33;
Best Local Similarity 100.0%; Pred. No. 5.9e-20;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRANONSMETRPDINPAMVYTGIRPVGRF 31
Db 1 SRANONSMETRPDINPAMVYTGIRPVGRF 31

RESULT 6
US-09-105-678A-7
Sequence 7, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-7

Query Match 93.7%; Score 163; DB 3; Length 31;
Best Local Similarity 93.5%; Pred. No. 2.6e-18;
Matches 29; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 SRANONSMETRPDINPAMVYTGIRPVGRF 31
Db 1 SRANONSMETRPDINPAMVYTGIRPVGRF 31

RESULT 7
US-09-105-678A-31
Sequence 31, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:

LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-31

Query Match 93.7%; Score 163; DB 3; Length 31;
Best Local Similarity 93.5%; Pred. No. 2,6e-18;
Matches 29; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SRAHSHMETRPDINPAMWYGRGIRPVGRF 31
||||| ||||||| ||||||| |||||||
DB 1 SRAHSHMETRPDINPAMWYGRGIRPVGRF 31

RESULT 8

US-09-105-678A-32
Sequence 32, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 32 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-32

Query Match 93.7%; Score 163; DB 3; Length 32;
Best Local Similarity 93.5%; Pred. No. 2,7e-18;
Matches 29; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SRAHSHMETRPDINPAMWYGRGIRPVGRF 31
||||| ||||||| ||||||| |||||||
DB 1 SRAHSHMETRPDINPAMWYGRGIRPVGRF 31

RESULT 9

US-09-105-678A-33
Sequence 33, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-33

Query Match 93.7%; Score 163; DB 3; Length 33;
Best Local Similarity 93.5%; Pred. No. 2,8e-18;
Matches 29; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SRAHSHMETRPDINPAMWYGRGIRPVGRF 31
||||| ||||||| ||||||| |||||||
DB 1 SRAHSHMETRPDINPAMWYGRGIRPVGRF 31

RESULT 10

US-09-105-678A-29
Sequence 29, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 29 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-29

Query Match 87.4%; Score 152; DB 3; Length 29;
Best Local Similarity 93.1%; Pred. No. 1,1e-16;
Matches 27; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SRAHSHMETRTPDINPAMTGRGIRPVG 29
DB 1 SRAHSHMETRTPDINPAMTGRGIRPVG 29

RESULT 11
US-09-105-678A-9
Sequence 9, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-9

Query Match 85.6%; Score 149; DB 3; Length 31;
Best Local Similarity 83.9%; Pred. No. 3.4e-16;
Matches 26; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 SRAHSHMETRTPDINPAMTGRGIRPVGR 31
DB 1 SRAHSHMETRTPDINPAMTGRGIRPVGR 31

RESULT 12
US-09-105-678A-43
Sequence 43, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-43

Query Match 85.6%; Score 149; DB 3; Length 31;
Best Local Similarity 83.9%; Pred. No. 3.4e-16;
Matches 26; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 SRAHSHMETRTPDINPAMTGRGIRPVGR 31
DB 1 SRAHSHMETRTPDINPAMTGRGIRPVGR 31

RESULT 13
US-09-105-678A-44
Sequence 44, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
CITY: Boston
STREET: 130 Water Street
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 32 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-44
Query Match 85.6%; Score 149; DB 3; Length 32;
Best Local Similarity 83.9%; Pred. No. 3.6e-16;
Matches 26; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 1 SRAHNSMETRTDINPAMTGRGIRPVGRF 31
DB 1 SRTHRSMETRTDINPAMTASRGIRPVGRF 31
RESULT 14
US-09-105-678A-45
Sequence 45, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
CITY: Boston
STREET: 130 Water Street
STATE: MA
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-45
Query Match 85.6%; Score 149; DB 3; Length 33;
Best Local Similarity 83.9%; Pred. No. 3.7e-16;
Matches 26; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 1 SRAHNSMETRTDINPAMTGRGIRPVGRF 31
DB 1 SRTHRSMETRTDINPAMTASRGIRPVGRF 31
RESULT 15
US-09-105-678A-40
Sequence 40, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
CITY: Boston
STREET: 130 Water Street
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-40

Query Match 66.7%; Score 116; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.1e-11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 12 TPDIAPAWYTGRIKRVGRF 31
|||||
Db 1 TPDIAPAWYTGRIKRVGRF 20

Search completed: April 17, 2001, 15:39:48
Job time: 317 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 17, 2001, 15:38:42 ; Search time 116.94 Seconds
(without alignments)
15.154 Million cell updates/sec

Title: US-09-446-543a-47
Perfect score: 174
Sequence: 1 SRAHSHMETRPDINPANTYGRGIRPGRF 31

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A.Geneseq_0401.*
1: /SIDS1/gcgdata/geneseq/geneseq/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseq/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseq/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseq/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseq/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseq/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseq/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseq/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseq/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseq/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseq/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseq/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseq/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseq/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseq/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseq/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseq/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseq/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseq/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	174	100.0	31	W31384	Rat type G protein
2	174	100.0	31	W97233	Rat type G protein
3	174	100.0	31	W87614	Rat type G protein
4	174	100.0	31	W95173	Murine pituitary-d
5	174	100.0	31	W95174	Murine pituitary-d
6	174	100.0	31	B10355	Rat oxytocin secre
7	174	100.0	31	Y49294	Rat prolactin-rele
8	174	100.0	31	Y49292	19P2 ligand peptid
9	174	100.0	32	W31385	Rat type G protein
10	174	100.0	32	B10356	Rat oxytocin secre
11	174	100.0	33	W31386	Rat type G protein

12	174	100.0	33	21	B10357	Rat oxytocin secre
13	174	100.0	82	20	W95172	Murine pituitary-d
14	174	100.0	83	18	W31383	Rat type G protein
15	174	100.0	83	20	W97225	Rat type G protein
16	174	100.0	83	21	B10354	Rat oxytocin secre
17	163	93.7	31	18	W31371	Bovine G protein-c
18	163	93.7	31	20	W97218	Bovine pituitary-d
19	163	93.7	31	20	W87613	Bovine 19P2 ligand
20	163	93.7	31	20	W95188	Bovine pituitary-d
21	163	93.7	31	21	B10347	Bovine oxytocin se
22	163	93.7	31	21	Y49290	19P2 ligand peptid
23	163	93.7	31	21	Y49298	19P2 ligand peptid
24	163	93.7	32	18	W31372	Bovine G protein-c
25	163	93.7	32	20	W95189	Bovine pituitary-d
26	163	93.7	32	21	B10348	Bovine oxytocin se
27	163	93.7	33	18	W31373	Bovine G protein-c
28	163	93.7	33	20	W95190	Bovine pituitary-d
29	163	93.7	33	21	B10349	Bovine oxytocin se
30	163	93.7	33	21	Y49297	19P2 ligand peptid
31	163	93.7	98	18	W31382	Bovine genome deri
32	163	93.7	98	18	W31368	Bovine G protein-c
33	163	93.7	98	20	W97224	Bovine genome-driv
34	163	93.7	98	20	W97217	Bovine pituitary-d
35	163	93.7	98	20	W95187	Bovine genome-deri
36	163	93.7	98	21	B10346	Bovine oxytocin se
37	163	93.7	98	21	B10353	Bovine G protein-c
38	152	87.4	29	18	W31369	Bovine G protein-c
39	152	87.4	29	20	W95184	Bovine pituitary-d
40	149	85.6	31	18	W31391	Human type G prote
41	149	85.6	31	20	W97235	Human type G prote
42	149	85.6	31	20	W87615	Human 19P2 ligand
43	149	85.6	31	21	B10362	Human oxytocin sec
44	149	85.6	31	21	Y49291	19P2 ligand peptid
45	149	85.6	32	18	W31392	Human type G prote

ALIGNMENTS

RESULT 1	W31384	W31384 standard; Peptide: 31 AA.
XX	AC	W31384:
XX	DF	06-APR-1998 (first entry)
XX	XX	Rat type G protein-coupled receptor ligand fragment 1.
XX	DE	G protein-coupled receptor; ligand binding; pharmaceutical;
XX	KW	modulator; pituitary; central nervous system; pancreas; prophylactic;
KW	KW	therapeutic agent.
XX	XX	Rat sp.
OS	XX	W09724436-A2.
XX	PN	10-JUL-1997.
XX	PD	26-DEC-1996; 96WO-UF03821.
XX	XX	18-SEP-1996; 96JP-0246573.
PR	PR	28-DEC-1995; 95JP-0343371.
PR	PR	15-MAR-1996; 96JP-0059419.
PR	PR	12-AUG-1996; 96JP-0211805.
XX	XX	(TAKE) TAKEDA CHEM IND LTD.
PA	XX	Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;
XX	PI	Kawamata Y, Kitada C;
XX	XX	WPI; 1997-363672/33.
DR	DR	N-FSDB; V02421.

XX Ligand peptide for G protein-coupled receptor - acts by modulating
 PT function in the central nervous system, pancreas and pituitary gland
 XX
 PS Claim 2; Page 179; 258pp; English.
 CC This sequence represents a peptide fragment from a novel rat type
 CC ligand polypeptide corresponding to amino acid residues 22 to 52 of the
 CC sequence represented in W31383 and is used in an assay to monitor ligand
 CC binding to the G protein-coupled receptor protein. Pharmaceutical
 CC compositions containing this ligand may be used as a pituitary function
 CC modulator, a central nervous system modulator or a pancreatic function
 CC modulator. This ligand could have specific applications as a prophylactic
 CC or therapeutic agent for dementia, depression, hyperkinetic syndrome,
 CC disturbance of consciousness, anxiety syndrome, schizophrenia, trauma,
 CC growth hormone secretory disease, hyper- and polyphagia, hyperlipidaemia,
 CC hypercholesterolemia, hyperglycaemia, hyperprolactinaemia, diabetes,
 CC cancer, pancreatitis, renal disease, Turner's syndrome, neurosis, asthma,
 CC rheumatoid arthritis, spinal injury, transient brain ischaemia, epilepsy,
 CC amyotrophic lateral sclerosis, acute myocardial infarction, infertility,
 CC splenomegaly, degeneration, bone fracture, trauma, atopic dermatitis,
 CC osteoporosis and/or oligodactylia. Assays can also be developed to screen
 CC compounds which are capable of altering the binding activity of the
 CC ligand affecting activation of the G protein-coupled receptor protein.
 CC
 SQ Sequence 31 AA;

Query Match 100.0%; Score 174; DB 16; Length 31;
 Best Local Similarity 100.0%; Pred. No. 5.8e-19;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRAHSHMETRPDIPNAPWYTGIRPVGRF 31
 ||||||||||||||||||||||||||||
 DB 1 srahshmetrpdipnawytgrtgrpvgrf 31

RESULT 2
 W97233
 ID W97233 standard; peptide; 31 AA.

AC W97233;

DT 06-MAY-1999 (first entry)

DE Rat type ligand polypeptide fragment.

XX Rat type ligand; modulation; prolactin secretion;
 KW G protein-coupled receptor; GPCR; hypovarianism; gonocyst cacosgenesis;
 KW menopausal syndrome; euthyroid; hypometabolism; lactation;
 KW pituitary adenomatosis; brain tumour; emmenopathy; autoimmune disease;
 KW prolactinoma; infertility; impotence; amenorrhea; galactorrhea;
 KW acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;
 KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome or dyszoospermia;
 KW contraceptive; placental function; chorioncarcinoma; hydatid mole;
 KW interruption mole; abortion; unfertilized fetus; abnormal saccharometabolism;
 KW abnormal lipidmetabolism; oxytocia.

OS Rattus sp.

XX W09858962-A1.

PD 30-DEC-1998.

PF 22-JUN-1998; 98WO-JP02765.

PR 23-JUN-1997; 97JP-0165437.

PA (TAKE) TAKEDA CHEM IND LTD.

XX Fujii R, Hinuma S, Kawamata Y, Matsumoto H;

XX WPI; 1999-105614/09.

XX Use of G protein-coupled receptor ligands - for modulating prolactin
 PT secretion or placental function, e.g. for treating menopausal
 PT syndrome, tumours, autoimmune disease or abnormal pregnancy
 XX
 PS Claim 3; Page 153; 241pp; English.

CC The present sequence represents a rat type ligand fragment. It
 CC is used in the course of the invention. The specification describes
 CC an agent for modulating prolactin secretion which comprises a
 CC ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)
 CC protein. The agents for promoting prolactin secretion can be used for
 CC treating or preventing hypovarianism, gonocyst cacosgenesis, menopausal
 CC syndrome, euthyroid or hypometabolism. They can be used for promoting
 CC lactation in a domestic mammal and as an aphrodisiac. The agents for
 CC inhibiting prolactin secretion can be used for treating or preventing
 CC pituitary adenomatosis, brain tumour, emmenopathy, autoimmune disease,
 CC prolactinoma, infertility, impotence, amenorrhea, galactorrhea,
 CC acromegaly, Chiari-Frommel syndrome, Argonz-del Castillo syndrome,
 CC Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dyszoospermia.
 CC The inhibitory agents can also be used as contraceptives. The agents for
 CC modulating placental function can be used for treating or preventing
 CC chorioncarcinoma, hydatid mole, abortion, unfertilized fetus,
 CC abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.
 CC
 SQ Sequence 31 AA;

Query Match 100.0%; Score 174; DB 20; Length 31;
 Best Local Similarity 100.0%; Pred. No. 5.8e-19;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRAHSHMETRPDIPNAPWYTGIRPVGRF 31
 ||||||||||||||||||||||||||||
 DB 1 srahshmetrpdipnawytgrtgrpvgrf 31

RESULT 3
 W87614
 ID W87614 standard; Peptide; 31 AA.

AC W87614;

DT 29-MAR-1999 (first entry)

DE Rat 19P2 ligand.

XX 19P2 ligand; G protein coupled receptor; pituitary;
 KW prolactin releasing peptide; rat; dementia; breast cancer;
 KW therapy.

OS Rattus sp.

PN EP087417-A2.

PD 30-DEC-1998.

PF 25-JUN-1998; 98EP-0111725.

PR 27-JUN-1997; 97JP-0172118.

PA (TAKE) TAKEDA CHEM IND LTD.

XX Moriya T, Nishimura O, Suenaga M, Tanaka Y;

XX WPI; 1999-047884/05.

XX Producing a 19P2 pituitary G protein receptor ligand - by cleavage
 PT of a fusion protein, useful for preventing and treating dementia,
 PT breast cancer, renal failure and autoimmune disease
 XX
 PS Claim 5; Page 34; 56pp; English.

CC This is the amino acid sequence of the rat pituitary G
 CC protein-coupled receptor ligand 19P2L. A method suitable for
 CC commercial high-level production of 19P2L comprises expressing
 CC the ligand in host cells as a recombinant fusion protein e.g. with
 CC human basic fibroblast growth factor (see V83794-95) that has
 CC been modified to include an N-terminal cysteine residue. The
 CC ligand is released from the fusion by cyanylation followed by
 CC ammonolysis. 19P2L has prolactin secretion-stimulating and (at
 CC high doses) prolactin secretion-inhibiting properties. It can be
 CC used in the treatment and prevention of various diseases including:
 CC senile dementia, cerebrovascular dementia, and dementia associated
 CC with: neurological disorders (e.g. Alzheimer's disease, Parkinson's
 CC disease, Pick's disease, Huntington's disease), infectious diseases
 CC (e.g. Creutzfeldt-Jakob's), endocrine or metabolic disease or
 CC toxicosis (e.g. hypothyroidism, vitamin B12 deficiency, alcoholism,
 CC intoxication by drugs, metal and organic compounds), tumorigenic
 CC diseases (e.g. brain tumour), traumatic diseases (e.g. chronic
 CC subarachnoidal haemorrhage, and other types of dementia, depression,
 CC hyperactive child syndrome (microencephalopathy) and disturbance of
 CC consciousness. It is also useful for prevention and treatment of
 CC diseases associated with prolactin hypo and hypersecretion
 CC respectively, including: hyperprolactinemia, pituitary adenoma,
 CC breast cancer, infertility, impotence and autoimmune disease
 CC (hypersecretion disorders), and seminal vesicle hypoplasia,
 CC osteoporosis, menopausal syndrome and renal failure (hyposecretion
 CC disorders). The 19P2 polypeptide/amide is also useful as a test
 CC reagent for study of the prolactin secretory function or as a
 CC lactagogue in mammalian farm animals.

SO Sequence 31 AA:

Query Match 100.0%; Score 174; DB 20; Length 31;
 Best Local Similarity 100.0%; Pred. No. 5.8e-19;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRAHOSMETRPDINPAWYTGRIPIVGRF 31
 |||
 DB 1 strahgsmetrtppdinpawyltgrlrvgrf 31

RESULT 4
 W95173
 ID W95173 standard; peptide; 31 AA.

AC W95173;

DT 10-MAR-1999 (first entry)

DE Murine pituitary-derived ligand mature polypeptide sequence.

KM Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;
 KM GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas;
 KM tissue; screen; therapeutic; binding; senile dementia; ligand; murine;
 KM Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;
 KM Creutzfeld-Jakob disease; poisoning; schizophrenia; growth hormone;
 KM secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;
 KM gene therapy; transgenic animal.

KW Mus sp.

OS W09849295-A1.

PN 05-NOV-1998.

PD 27-APR-1998; 98WO-JP01923.

PF 28-APR-1997; 97JP-0109974.

PR (TAKE) TAKEDA CHEM IND LTD.

PA Fukusumi S, Hinuma S;

XX

DR WPI; 1999-009423/01.
 XX New polypeptide ligand for orphan G protein coupled receptors - used
 PT for treating disorders of central nervous system, pituitary and
 PT pancreas, and for drug screening
 XX
 PS Disclosure; Page 134; 206pp; English.

CC This represents the matured murine pituitary-derived ligand polypeptide
 CC sequence. The polypeptide is a ligand for the G-protein coupled orphan
 CC receptor designated GPR10 (human) or UHR-1 (rat). Cells transformed with
 CC a vector containing the ligand polypeptide encoding DNA are used to
 CC produce a recombinant ligand polypeptide. The ligand polypeptide, and its
 CC fragments, modulate function of the pituitary, central nervous system,
 CC pancreas and other tissues and can be used to screen for agents that
 CC modulate binding of the polypeptide to the receptor; to quantify the
 CC amount of receptor in a sample and to raise antibodies. They may also be
 CC used therapeutically, e.g. to treat senile dementia; Alzheimer's,
 CC Parkinson's or Huntington's diseases; Creutzfeld-Jakob disease; poisoning
 CC by heavy metals or drugs; diabetes; schizophrenia; disorders of growth
 CC hormone secretion; cancer; rheumatoid arthritis; epilepsy and many
 CC others, also to improve post-operative nutritional status and as
 CC vasopressor. Transgenic animals carrying the ligand polypeptide encoding
 CC DNA or its mutins are used to study the function of the polypeptide-
 CC expressing genes, as models of disease, for drug screening and as source
 CC of cell lines. The ligand polypeptide DNA is used as a source of probes
 CC and primers; to identify related sequences; in receptor-binding assays;
 CC for production of Ab and antisera; in drug development; for gene therapy
 CC and to develop transgenic animals.

SO Sequence 31 AA:

Query Match 100.0%; Score 174; DB 20; Length 31;
 Best Local Similarity 100.0%; Pred. No. 5.8e-19;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRAHOSMETRPDINPAWYTGRIPIVGRF 31
 |||
 DB 1 strahgsmetrtppdinpawyltgrlrvgrf 31

RESULT 5
 W95174
 ID W95174 standard; Protein; 31 AA.

AC W95174;

DT 10-MAR-1999 (first entry)

DE Murine pituitary-derived ligand polypeptide antigenic epitope.

KM Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;
 KM GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas;
 KM tissue; screen; therapeutic; binding; senile dementia; ligand; murine;
 KM Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;
 KM Creutzfeld-Jakob disease; poisoning; schizophrenia; growth hormone;
 KM secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;
 KM gene therapy; transgenic animal; epitope.

KW Mus sp.

OS W09849295-A1.

PN 05-NOV-1998.

PD 27-APR-1998; 98WO-JP01923.

PF 28-APR-1997; 97JP-0109974.

PR (TAKE) TAKEDA CHEM IND LTD.

PA Fukusumi S, Hinuma S;

XX

XX WPI; 1999-009423/01.

XX DR
XX PT New polypeptide ligand for orphan G protein coupled receptors - used
XX PT for treating disorders of central nervous system, pituitary and
XX PT pancreas, and for drug screening

XX PS Disclosure; Page 26; 206pp; English.

XX The invention relates to a murine pituitary-derived ligand polypeptide
XX which is a ligand for the G-protein coupled orphan receptor designated
XX GPR10 (human) or UHR-1 (rat). Cells transformed with a vector containing
XX the ligand polypeptide encoding DNA are used to produce a recombinant
XX ligand polypeptide. The ligand polypeptide, and its fragments, modulate
XX function of the pituitary, central nervous system, pancreas and other
XX tissues and can be used to screen for agents that modulate binding of the
XX polypeptide to the receptor; to quantify the amount of receptor in a
XX sample and to raise antibodies. They may also be used therapeutically,
XX e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's
XX diseases; Creutzfeldt-Jakob disease; poisoning by heavy metals or drugs;
XX diabetes; schizophrenia; disorders of growth hormone secretion; cancer;
XX rheumatoid arthritis, epilepsy and many others, also to improve post-
XX operative nutritional status and as vasopressor. Transgenic animals
XX carrying the ligand polypeptide encoding DNA or its mutagen are used to
XX study the function of the polypeptide-expressing genes, as models of
XX disease, for drug screening and as source of cell lines. The ligand
XX polypeptide DNA is used as a source of probes and primers; to identify
XX related sequences; in receptor-binding assays; for production of Ab and
XX antisera; in drug development; for gene therapy and to develop transgenic
XX animals. Sequences W95174 to W95178 represent antigenic epitopes which
XX can be used for the preparation of anti-ligand polypeptide antibody.

SO Sequence 31 AA;

Query Match 100.0%; Score 174; DB 20; Length 31;
Best Local Similarity 100.0%; Pred. No. 5.8e-19;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRAHSHMETRPDINPAWYTGRIYGVGRF 31
DB 1 SRAHSHMETRPDINPAWYTGRIYGVGRF 31

RESULT 6
B10355
ID B10355 standard; peptide; 31 AA.

XX AC B10355;

XX DT 24-NOV-2000 (first entry)

XX DE Rat oxytocin secretion promoting peptide SEQ ID NO: 18.

XX KW Rat; oxytocin secretion promoter; G protein-coupled receptor protein;
XX treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;
XX caesarean section; artificial fertilization; galactostasis; goat; pig;
XX veterinary medicine; milk production.

XX OS Rattus sp.

XX PN WO200038704-A1.

XX PD 06-JUL-2000.

XX PF 22-DEC-1999; 99WO-JP07199.

XX PR 25-DEC-1998; 98JP-0369585.

XX PA (TAKE) TAKEDA CHEM IND LTD.

XX PI Matsumoto H, Kitada C, Hinuma S;

DR WPI; 2000-452298/39.

XX DR
XX PT Physiologically-active polypeptide recognized as ligand by G
XX PT protein-coupled receptor protein, for promoting secretion of oxytocin,
XX PT as drugs for diseases relating to oxytocin secretion and in veterinary
XX PT medicine

XX PS Claim 3; Page 57; 72pp; Japanese.

XX This invention describes a novel oxytocin secretion-regulating agent
XX which contains a ligand peptide or its salt for the G-protein-coupled
XX receptor protein. It is useful in the form of drugs for ameliorating,
XX preventing and treating diseases relating to oxytocin secretion e.g.,
XX weak pains and atonic bleeding, before and after expulsion of placenta,
XX uterine recovery failure, caesarean section, stoppage of artificial
XX fertilization or galactostasis and is also applicable in veterinary
XX medicine for promoting milk production in cow, goat and pig. This
XX sequence represents a rat peptide which acts as an oxytocin secretion
XX promoter.

SO Sequence 31 AA;

Query Match 100.0%; Score 174; DB 21; Length 31;
Best Local Similarity 100.0%; Pred. No. 5.8e-19;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRAHSHMETRPDINPAWYTGRIYGVGRF 31
DB 1 SRAHSHMETRPDINPAWYTGRIYGVGRF 31

RESULT 7

Y87504
ID Y87504 standard; protein; 31 AA.

XX AC Y87504;

XX DT 18-JUL-2000 (first entry)

XX DE Rat prolactin-releasing peptide, PRRP.

XX KW Prolactin-releasing peptide; PRRP; GPR10; G protein-coupled receptor;
XX feeding behaviour; food intake; modulation; antagonist; anorectic;
XX obesity; agonist; cachexia.

XX OS Rattus sp.

XX FH Key Location/Qualifiers

XX FT Modified-site 31 /note="C-terminal amide"

XX PN WO200017641-A1.

XX PD 30-MAR-2000.

XX PF 22-SEP-1999; 99WO-US21243.

XX PR 22-SEP-1998; 98US-0101380.

XX PR 14-OCT-1998; 98US-0172353.

XX PA (MILL-) MILLENNIUM PHARM INC.

XX PI Stricker-Kongrad A, Gu W;

XX DR WPI; 2000-303231/26.

XX Identifying modulators of body weight by a combination of a cell-free
XX PT or cell-based assay to identify modulators of GPR10, followed by an in
XX vivo assay for the compounds effect on e.g. feeding behavior -
XX Example 2; Page 61; 82pp; English.

CC The invention relates to a method for identifying compounds useful for
 CC modulating body weight. The method comprises cell-free and/or cell-based
 CC assays that identify compounds which bind to and/or activate or inhibit
 CC the activity of GPR10, a G protein-coupled receptor. These assays are
 CC then followed by an in vivo assay of the effect of the compound on
 CC feeding behaviour, body weight or metabolic rate in a mammal. Proactin-
 CC releasing peptide (PRRP; Y87504) is a ligand of GPR10. Binding of PRRP to
 CC GPR10 stimulates a signal transduction cascade, which results in an
 CC increase in food intake. Compounds identified using the method of the
 CC invention are useful for the modulation of body weight. Antagonists of
 CC GPR10 can be used to treat obesity, while GPR10 agonists can be used to
 CC treat cachexia. The present sequence represents rat PRRP.

CC Sequence 31 AA:

Query Match 100.0%; Score 174; DB 21; Length 31;
 Best Local Similarity 100.0%; Pred. No. 5,8e-19;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHSHMETRPDINPAMYTGRIPVGRF 31
 DB 1 srahshmetrpdinpawtygrtprvgrf 31

RESULT 8

Y49292 Y49292 standard; peptide; 31 AA.

Y49292;

22-FEB-2000 (first entry)

19P2 ligand peptide fragment.

Monoclonal antibody; 19P2 ligand; diagnosis; prolactin secretion;
 pituitary; regulatory mechanism; central nervous system; pancreatic.

Rattus sp.

Key Location/Qualifiers
 Modified-site 31
 FT /note="C-terminal amide"

W09960112-AL.

25-NOV-1999.

20-MAY-1999; 99WO-JP02650.

21-MAY-1998; 98JP-0140293.

(TAKE) TAKEDA CHEM IND LTD.

Matsumoto H, Kitada C, Hinuma S;

WPI, 2000-039381/03.

New monoclonal antibodies, useful in diagnosis, as drugs and in
 studying diseases related to ligand abnormality.

Disclosure: Page 26; 73pp; Japanese.

The invention provides a monoclonal antibody which has a specific
 reaction with the part peptide of the C-terminal of 19P2 ligand or its
 derivative. The antibodies can be used in diagnosis or to treat or
 prevent diseases associated with abnormality in the pituitary function
 regulatory mechanism (e.g. promotion of prolactin secretion), central
 nervous regulatory mechanism, and pancreatic function regulatory
 mechanism. The antibody-based immunoassay can also be applied in
 clarifying the physiological functions of the ligand and its derivative.
 Sequences Y49290-302 represent peptide fragments of the 19P2 ligand.

SQ Sequence 31 AA:

Query Match 100.0%; Score 174; DB 21; Length 31;
 Best Local Similarity 100.0%; Pred. No. 5,8e-19;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRAHSHMETRPDINPAMYTGRIPVGRF 31
 DB 1 srahshmetrpdinpawtygrtprvgrf 31

RESULT 9

W31385 W31385 standard; peptide; 32 AA.

W31385;

06-APR-1998 (first entry)

Rat type G protein-coupled receptor ligand fragment 2.

G protein-coupled receptor; ligand binding; pharmaceutical;
 modulator; pituitary; central nervous system; pancreas; prophylactic;
 therapeutic agent.

Rat sp.

W09724436-A2.

10-JUL-1997.

26-DEC-1996; 96WO-JP03821.

18-SEP-1996; 96JP-0246573.

28-DEC-1995; 95JP-0343371.

15-MAR-1996; 96JP-0059419.

12-AUG-1996; 96JP-0211805.

(TAKE) TAKEDA CHEM IND LTD.

Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;

Kawamata Y, Kitada C;

WPI: 1997-363672/33.

N-PSDB: Y02422.

Ligand peptide for G protein-coupled receptor - acts by modulating
 function in the central nervous system, pancreas and pituitary gland
 Claim 2; Page 179; 258pp; English.

This sequence represents a peptide fragment from a novel rat type
 ligand polypeptide corresponding to amino acid residues 22 to 53 of the
 sequence represented in W31383 and is used in an assay to monitor ligand
 binding to the G protein-coupled receptor protein. Pharmaceutical
 compositions containing this ligand may be used as a pituitary function
 modulator, a central nervous system modulator or a pancreatic function
 modulator. This ligand could have specific applications as a prophylactic
 or therapeutic agent for dementia, depression, hyperkinetic syndrome,
 disturbance of consciousness, anxiety syndrome, schizophrenia, trauma,
 growth hormone secretory disease, hyper- and hypophagia, hyperlipidaemia,
 hypercholesterolaemia, hyperglycaemia, hyperprolactinaemia, diabetes,
 cancer, pancreatitis, renal disease, Turner's syndrome, neurosis, asthma,
 rheumatoid arthritis, spinal injury, transient brain ischaemia, epilepsy,
 amyotrophic lateral sclerosis, acute myocardial infarction, infertility,
 osteoporosis and/or oligosaccharia. Assays can also be developed to screen
 compounds which are capable of altering the binding activity of the
 ligand affecting activation of the G protein-coupled receptor protein.

Sequence 32 AA:

Query Match 100.0%; Score 174; DB 18; Length 32;
Best Local Similarity 100.0%; Pred. No. 6e-19;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRANQSMETRPDINPAMYTGKIRPVGRF 31
|||||
1 strahqsmetrrpdinpawytgrirpvgrrf 31

RESULT 10

ID B10356 standard; peptide: 32 AA.

AC B10356;

DE 24-NOV-2000 (first entry)

Rat oxytocin secretion promoting peptide SEQ ID NO: 19.

KM Rat; oxytocin secretion promoter; G protein-coupled receptor protein;
treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;
KM caesarean section; artificial fertilization; galactostasis; goat; pig;
KM veterinary medicine; milk production.

OS Rattus sp.

PN WO200038704-A1.

PD 06-JUL-2000.

PF 22-DEC-1999; 99WO-JP071199.

PR 25-DEC-1998; 98JP-0369585.

PA (TAKE) TAKEDA CHEM IND LTD.

PI Matsumoto H, Kitada C, Hinuma S;

DR WPI; 2000-452298/39.

PT Physiologically-active polypeptide recognized as ligand by G
protein-coupled receptor protein, for promoting secretion of oxytocin,
as drugs for diseases relating to oxytocin secretion and in veterinary
medicine -
PS Disclosure; Page 57; 72pp; Japanese.

CC This invention describes a novel oxytocin secretion-regulating agent
which contains a ligand peptide or its salt for the G protein-coupled
receptor protein. It is useful in the form of drugs for ameliorating,
preventing and treating diseases relating to oxytocin secretion e.g.,
weak pains and atonic bleeding, before and after expulsion of placenta,
uterine recovery failure, caesarean section, stoppage of artificial
fertilization or galactostasis and is also applicable in veterinary
medicine for promoting milk production in cow, goat and pig. This
sequence represents a rat peptide which acts as an oxytocin secretion
promoter.

CC XX
CC XX
SQ Sequence 32 AA;

Query Match 100.0%; Score 174; DB 21; Length 32;

Best Local Similarity 100.0%; Pred. No. 6e-19;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRANQSMETRPDINPAMYTGKIRPVGRF 31
|||||
1 strahqsmetrrpdinpawytgrirpvgrrf 31

Db 1 strahqsmetrrpdinpawytgrirpvgrrf 31

RESULT 11

ID W31386

ID W31386 standard; peptide: 33 AA.

AC W31386;

DE 06-APR-1998 (first entry)

Rat type G protein-coupled receptor ligand fragment 3.

KM G protein-coupled receptor; ligand binding; pharmaceutical;
modulator; pituitary; central nervous system; pancreas; prophylactic;
therapeutic agent.

OS Rat sp.

PN WO9724436-A2.

PD 10-JUL-1997.

PF 26-DEC-1996; 96WO-JP03821.

PR 18-SEP-1996; 96JP-0246573.

PR 28-DEC-1995; 95JP-0343371.

PR 15-MAR-1996; 96JP-0059419.

PR 12-AUG-1996; 96JP-0211805.

PA (TAKE) TAKEDA CHEM IND LTD.

PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;

PI Kawamata Y, Kitada C;

DR WPI; 1997-363672/33.

DR N-PSDB; V02423.

PT Ligand peptide for G protein-coupled receptor - acts by modulating
function in the central nervous system, pancreas and pituitary gland

PS Claim 2; Page 179-180; 258pp; English.
CC This sequence represents a peptide fragment from a novel rat type
ligand polypeptide corresponding to amino acid residues 22 to 54 of the
sequence represented in W31383 and is used in an assay to monitor ligand
binding to the G protein-coupled receptor protein. Pharmaceutical
compositions containing this ligand may be used as a pituitary function
modulator, a central nervous system modulator or a pancreatic function
modulator. This ligand could have specific applications as a prophylactic
or therapeutic agent for dementia, depression, hyperkinetic syndrome,
disturbance of consciousness, anxiety syndrome, schizophrenia, trauma,
growth hormone secretory disease, hyper- and polyphagia, hyperlipidaemia,
hypercholesterolaemia, hyperglycaemia, hyperprolactinaemia, diabetes,
cancer, pancreatitis, renal disease, Turner's syndrome, neurosis, asthma,
rheumatoid arthritis, spinal injury, transient brain ischaemia, epilepsy,
CC amyotrophic lateral sclerosis, acute myocardial infarction, infertility,
CC spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis,
CC osteoporosis and/or oligosaccharia. Assays can also be developed to screen
CC compounds which are capable of altering the binding activity of the
ligand affecting activation of the G protein-coupled receptor protein.

CC XX
CC XX
SQ Sequence 33 AA;

Query Match 100.0%; Score 174; DB 18; Length 33;

Best Local Similarity 100.0%; Pred. No. 6.2e-19;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRANQSMETRPDINPAMYTGKIRPVGRF 31
|||||
1 strahqsmetrrpdinpawytgrirpvgrrf 31

Db 1 strahqsmetrrpdinpawytgrirpvgrrf 31

RESULT 12

ID B10357 standard; peptide: 33 AA.

ID W31386

```

AC B10357;
XX
XX 24-NOV-2000 (first entry)
DT
XX
XX Rat oxytocin secretion promoting peptide SEQ ID NO: 20.
DE
XX
XX Rat; oxytocin secretion promoter; G-protein-coupled receptor protein;
KM treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;
KM caesarean section; artificial fertilization; galactostasis; goat; pig;
KM veterinary medicine; milk production.
XX
XX Rattus sp.
OS
XX WO200038704-A1.
XX
XX 06-JUL-2000.
XX
XX 22-DEC-1999; 59WO-JP07199.
XX
XX 25-DEC-1998; 98JP-0369585.
XX
XX (TAKE ) TAKEDA CHEM IND LTD.
XX
XX Matsumoto H, Kitada C, Hinuma S;
XX
XX WPI: 2000-452298/39.
XX
XX Physiologically-active polypeptide recognized as ligand by G
XX protein-coupled receptor protein, for promoting secretion of oxytocin,
XX as drugs for diseases relating to oxytocin secretion and in veterinary
XX medicine
XX
XX Disclosure: Page 58; 72pp; Japanese.
XX
XX This invention describes a novel oxytocin secretion-regulating agent
XX which contains a ligand peptide or its salt for the G-protein-coupled
XX receptor protein. It is useful in the form of drugs for ameliorating,
XX preventing and treating diseases relating to oxytocin secretion e.g.,
XX weak pains and atonic bleeding, before and after expulsion of placenta,
XX uterine recovery failure, caesarean section, stoppage of artificial
XX fertilization or galactostasis and is also applicable in veterinary
XX medicine for promoting milk production in cow, goat and pig. This
XX sequence represents a rat peptide which acts as an oxytocin secretion
XX promoter.
XX
XX Sequence 33 AA:
SQ
XX
XX Query Match 100.0%; Score 174; DB 21; Length 33;
XX Best Local Similarity 100.0%; Pred. No. 6,2e-19;
XX Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 SRAHQSMETRPDPINPAMYTGRIYVGRF 31
XX ||||||||||||||||||||||||||||
DB 1 stahghsmetrpdpinpawytgrgtrpygrf 31
XX
XX
XX RESULT 13
XX W95172
XX ID W95172 standard; Protein; 82 AA.
XX
XX AC W95172;
XX
XX 10-MAR-1999 (first entry)
XX
XX Murine pituitary-derived ligand polypeptide.
XX
XX Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;
XX GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas;
XX tissue; screen; therapeutic; binding; senile dementia; ligand; murine;
XX Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;
XX Creutzfeldt-Jakob disease; poisoning; schizophrenia; growth hormone;
XX secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;

```

```

KM gene therapy; transgenic animal.
XX
XX Mus sp.
XX
XX Key Location/Qualifiers
XX Peptide 1..20
XX Protein /note="signal peptide"
XX 21..82
XX /note="mature protein"
XX
XX WO9849295-A1.
XX
XX 05-NOV-1998.
XX
XX 27-APR-1998; 98WO-JP01923.
XX
XX 28-APR-1997; 97JP-0109974.
XX
XX (TAKE ) TAKEDA CHEM IND LTD.
XX
XX Fukusumi S, Hinuma S;
XX
XX WPI: 1999-009423/01.
XX
XX N-PSDB; V81228, V81229.
XX
XX New polypeptide ligand for orphan G protein coupled receptors - used
XX for treating disorders of central nervous system, pituitary and
XX pancreas, and for drug screening
XX
XX Claim 1; Page 133; 206pp; English.
XX
XX This represents a murine pituitary-derived ligand polypeptide which is a
XX ligand for the G-protein coupled orphan receptor designated GPR10 (human)
XX or UHR-1 (rat). Cells transformed with a vector containing the ligand
XX polypeptide encoding DNA are used to produce a recombinant ligand
XX polypeptide. The ligand polypeptide, and its fragments, modulate function
XX of the pituitary, central nervous system, pancreas and other tissues and
XX can be used to screen for agents that modulate binding of the polypeptide
XX to the receptor; to quantify the amount of receptor in a sample and to
XX raise antibodies. They may also be used therapeutically, e.g. to treat
XX senile dementia; Alzheimer's, Parkinson's or Huntington's diseases;
XX Creutzfeldt-Jakob disease; poisoning by heavy metals or drugs; diabetes;
XX schizophrenia; disorders of growth hormone secretion; cancer; rheumatoid
XX arthritis; epilepsy and many others, also to improve post-operative
XX nutritional status and as vasopressor. Transgenic animals carrying the
XX ligand polypeptide encoding DNA or its mutin are used to study the
XX function of the polypeptide-expressing genes, as models of disease, for
XX drug screening and as source of cell lines. The ligand polypeptide DNA is
XX used as a source of probes and primers; to identify related sequences; in
XX receptor-binding assays; for production of Ab and antisera; in drug
XX development; for gene therapy and to develop transgenic animals.
XX
XX Sequence 82 AA:
SQ
XX
XX Query Match 100.0%; Score 174; DB 20; Length 82;
XX Best Local Similarity 100.0%; Pred. No. 1.8e-18;
XX Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 SRAHQSMETRPDPINPAMYTGRIYVGRF 31
XX ||||||||||||||||||||||||||||
DB 21 stahghsmetrpdpinpawytgrgtrpygrf 51
XX
XX
XX RESULT 14
XX W31383
XX ID W31383 standard; Protein; 83 AA.
XX
XX AC W31383;
XX
XX 06-APR-1998 (first entry)
XX
XX Rat type G protein-coupled receptor ligand encoded by pRAV3.

```

XX G protein-coupled receptor; ligand binding; pharmaceutical;
 KW modulator; pituitary; central nervous system; pancreas; prophylactic;
 KW therapeutic agent.
 OS Rat sp.
 XX WO9724436-A2.
 XX 10-JUL-1997.
 XX 26-DEC-1996; 96WO-JP03821.
 XX 18-SEP-1996; 96JP-0246573.
 XX 28-DEC-1995; 95JP-0343371.
 XX 15-MAR-1996; 96JP-0059419.
 XX 12-AUG-1996; 96JP-0211805.
 XX (TAKE) TAKEDA CHEM IND LTD.
 XX Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M,
 PI Kawamata Y, Kitada C;
 XX WPI: 1997-363672/33.
 DR N-PSDB; V02420.
 XX Ligand peptide for G protein-coupled receptor - acts by modulating
 PT function in the central nervous system, pancreas and pituitary gland
 PT
 XX Claim 3; Page 178; 258pp; English.
 XX This sequence represents a novel rat type ligand polypeptide encoded by
 CC PRAV3 which is used in an assay to monitor ligand binding to the G
 CC protein-coupled receptor protein. Pharmaceutical compositions
 CC containing this ligand may be used as a pituitary function modulator, a
 CC central nervous system modulator or a pancreatic function modulator. This
 CC ligand could have specific applications as a prophylactic or therapeutic
 CC agent for dementia, depression, hyperkinetic syndrome, disturbance of
 CC consciousness, anxiety syndrome, schizophrenia, trauma, growth hormone
 CC secretory disease, hyper- and polyphagia, hypercholesterolaemia,
 CC hyperglycaemia, hyperlipidaemia, hyperprolactinaemia, diabetes,
 CC cancer, pancreatitis, renal disease, Turner's syndrome, neurosis,
 CC amyotrophic lateral sclerosis, spinal injury, transient brain ischaemia,
 CC spino cerebellar degeneration, bone fracture, trauma, atopic dermatitis,
 CC osteoporosis, asthma, epilepsy, infertility and/or oligosaccharia. Assays
 CC can also be developed to screen compounds which are capable of altering
 CC the binding activity of the ligand thus affecting activation of the G
 CC protein-coupled receptor protein.
 CC
 SQ Sequence 83 AA;
 Query Match 100.0%; Score 174; DB 18; Length 83;
 Best Local Similarity 100.0%; Pred. No. 1.8e-18;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SRHOSHMETRTDIPNPAWYTGIRPVGRF 31
 DB 22 strahqsmetrtppdinpawtgrgtrpvgrf 52
 RESULT 15
 W97225
 ID W97225 standard; Peptide; 83 AA.
 AC W97225;
 XX 06-MAY-1999 (first entry)
 DT
 XX Rat type ligand polypeptide.
 DE
 XX G protein-coupled receptor; GPCR; hypovarianism; gonocyst caecogenesis;
 KW

KW menopausal syndrome; euthyroid; hypometabolism; lactation; modulation;
 KW pituitary adenomatosis; brain tumour; emmenopathy; autoimmune disease;
 KW prolactinoma; infertility; impotence; amenorrhea; galactorrhea;
 KW acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;
 KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyscoospermia;
 KW contraceptive; placental function; choriocarcinoma; hydatid mole;
 KW interruption mole; abortion; unthrifty fetus; abnormal saccharometabolism;
 KW abnormal lipidmetabolism; oxytocia; prolactin secretion;
 KW rat type ligand.
 OS Rattus sp.
 XX WO9858962-A1.
 XX 30-DEC-1998.
 XX 22-JUN-1998; 98WO-JP02765.
 XX 23-JUN-1997; 97JP-0165437.
 XX (TAKE) TAKEDA CHEM IND LTD.
 XX Fujii R, Hinuma S, Kawamata Y, Matsumoto H;
 PI WPI: 1999-105614/09.
 DR N-PSDB; X15525.
 XX Use of G protein-coupled receptor ligands - for modulating prolactin
 PT secretion or placental function, e.g. for treating menopausal
 PT syndrome, tumours, autoimmune disease or abnormal pregnancy
 PT
 XX Disclosure; Page 152; 241pp; English.
 XX The present sequence represents a rat type ligand polypeptide. The
 CC specification describes an agent for modulating prolactin secretion
 CC which comprises a ligand polypeptide or a salt, for a G protein-coupled
 CC receptor (GPCR) protein. The agents for promoting prolactin secretion
 CC can be used for treating or preventing hypovarianism, gonocyst
 CC caecogenesis, menopausal syndrome, euthyroid or hypometabolism. They
 CC can be used for promoting lactation in a domestic mammal and as an
 CC aphrodisiac. The agents for inhibiting prolactin secretion can be used
 CC for treating or preventing pituitary adenomatosis, brain tumour,
 CC amenorrhoea, galactorrhea, acromegaly, Chiari-Frommel syndrome, Argonz-del
 CC Castillo syndrome, Forbes-Albright syndrome, lymphoma, Sheehan syndrome
 CC or dyscoospermia. The inhibitory agents can also be used as
 CC contraceptives. The agents for modulating placental function can be used
 CC for treating or preventing choriocarcinoma, hydatid mole, interruption mole,
 CC abortion, unthrifty fetus, abnormal saccharometabolism, abnormal
 CC lipidmetabolism or oxytocia.
 CC
 SQ Sequence 83 AA;
 Query Match 100.0%; Score 174; DB 20; Length 83;
 Best Local Similarity 100.0%; Pred. No. 1.8e-18;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SRHOSHMETRTDIPNPAWYTGIRPVGRF 31
 DB 22 strahqsmetrtppdinpawtgrgtrpvgrf 52
 Search completed: April 17, 2001, 15:38:42
 Job time: 349 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 17, 2001, 15:48:07 ; Search time 115.07 Seconds
(without alignments)
20.372 Million cell updates/sec

Title: US-09-446-543a-50
Perfect score: 116
Sequence: 1 TPDINPAMYTGRIKRPVGRF 20

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 374700 seqs, 117207915 residues
Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP_mhc:*
8: SP_organella:*
9: SP_phage:*
10: SP_plant:*
11: SP_protent:*
12: SP_unclassified:*
13: SP_vertebrate:*
14: SP_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	77.6	117	13	Q9W624
2	53	45.7	54	4	Q9UJF9
3	53	45.7	465	4	Q60687
4	49	42.2	419	4	Q9Y276
5	48	41.4	223	2	Q85605
6	48	41.4	790	10	Q9M371
7	47	40.5	430	1	Q27142
8	47	40.5	940	2	Q9KUM5
9	46.5	40.1	333	2	Q9RU10
10	46	39.7	333	2	Q9PH76
11	46	39.7	527	5	Q76383
12	46	39.7	555	3	Q00050
13	46	39.7	820	5	Q9NE93
14	45	38.8	267	5	Q9V3E5
15	45	38.8	330	4	Q9NRB7
16	45	38.8	767	5	Q20170
17	45	38.8	965	2	Q9PAR9
18	44	37.9	105	2	Q9PM69
19	44	37.9	306	2	Q9LAD4

20	44	37.9	320	2	Q9L7X4	Q9L7X4 pseudomonas
21	44	37.9	343	3	Q74569	Q74569 coprinus ci
22	44	37.9	398	10	Q9S7R5	Q9S7R5 arabidopsis
23	44	37.9	414	2	Q33480	Q33480 propionibac
24	44	37.9	418	2	Q9RT13	Q9RT13 delinococcus
25	44	37.9	449	5	Q01498	Q01498 caenorhabdi
26	44	37.9	540	10	Q9L6Z0	Q9L6Z0 oryza sativ
27	44	37.9	548	10	Q9L6Z0	Q9L6Z0 arabidopsis
28	44	37.9	775	1	Q9P9K4	Q9P9K4 pyrococcus
29	43.5	37.5	506	2	Q9ZAL1	Q9ZAL1 pseudomonas
30	43.5	37.5	1501	10	Q9SD86	Q9SD86 arabidopsis
31	43.5	37.5	1679	5	Q24301	Q24301 diosiphila
32	43	37.1	152	10	Q39562	Q39562 chlamydomon
33	43	37.1	309	5	Q17234	Q17234 caenorhabdi
34	43	37.1	429	2	Q9LOX9	Q9LOX9 streptomyces
35	43	37.1	498	10	Q9STP4	Q9STP4 arabidopsis
36	43	37.1	503	2	Q9KRT1	Q9KRT1 vibrio chol
37	43	37.1	517	10	Q9STP3	Q9STP3 arabidopsis
38	43	37.1	627	2	Q9L8G6	Q9L8G6 clostridium
39	43	37.1	630	4	Q9NKS8	Q9NKS8 homo sapien
40	43	37.1	948	2	Q9JUS4	Q9JUS4 neisseria m
41	43	37.1	949	2	Q9JUS4	Q9JUS4 neisseria m
42	43	37.1	1282	2	Q9Z6S5	Q9Z6S5 heliobacill
43	42.5	36.6	243	2	Q9REH5	Q9REH5 agrobacteri
44	42.5	36.6	443	5	Q19879	Q19879 caenorhabdi
45	42.5	36.6	506	2	Q9KH02	Q9KH02 azotobacter

ALIGNMENTS

RESULT 1
ID Q9W624 PRELIMINARY; PRT; 117 AA.
AC Q9W624;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE C-RF AMIDE PRECURSOR.
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Cyprininae; Carassius..
OX MCB1_taxid=7957;
FM [1]
RP SEQUENCE FROM M.A.
RC TISSUE-BRAIN;
RA Satake H., Minakata H., Fujimoto M.;
RT "Carassius Rhamde (C-RF amide).";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB020024; BAA76662.1; -; D5DC4CB22038C2B0 CRC64;
SQ SEQUENCE 117 AA; 12879 MW; D5DC4CB22038C2B0 CRC64;

Query Match 77.6%; Score 90; DB 13; Length 117;
Best Local Similarity 65.0%; Pred. No. 2.3e-06;
Matches 13; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 TPDINPAMYTGRIKRPVGRF 20
DB 56 SPEDIDPFWYGVGRVPIGRF 75

RESULT 2
ID Q9UJF9 PRELIMINARY; PRT; 54 AA.
AC Q9UJF9;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE D479J7.3 (SUSHI-REPEAT PROTEIN (SRPUL)) (FRAGMENT).
GN D479J7.3.
OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 RN NCBI_TaxID=9606;
 RP [1]
 RP SEQUENCE FROM N.A.
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL035608; CAB55682.1; -
 FT NON_TER 54
 SQ SEQUENCE 54 AA; 6110 MW; E2F3C39F7B961A9F CRC64;

Query Match 45.7%; Score 53; DB 4; Length 54;
 Best Local Similarity 50.0%; Pred. No. 0.41;
 Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 TPDINPMTGRCIRP 16
 II : ||| | | |
 DB 18 TPVPTMTAGSGYYP 33

RESULT 3
 ID 060687 PRELIMINARY; PRT; 465 AA.
 AC 060687;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE SUSHI-REPEAT PROTEIN.
 GN SRPUL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kurosawa H., Inukai T., Inaba T., Gol K., Chang K.-S., Shinjo T.,
 RA Rakestraw K.M., Naeye C.W., Look T.A.;
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF060567; AAC15765.1; -
 DR INTERPRO: IPR000436; -
 DR INTERPRO: IPR001128; -
 DR PPM: PF00084; sushi; 3.
 DR PROSITE: PS00086; CYTOCHROME P450; UNKNOWN.1.
 SQ SEQUENCE 465 AA; 52971 MW; 4D752B187F3E9FB8 CRC64;

Query Match 45.7%; Score 53; DB 4; Length 465;
 Best Local Similarity 50.0%; Pred. No. 4.3;
 Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 TPDINPMTGRCIRP 16
 II : ||| | | |
 DB 18 TPVPTMTAGSGYYP 33

RESULT 4
 ID 091276 PRELIMINARY; PRT; 419 AA.
 AC 091276;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE H-BCSL.
 GN BCSL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BRAIN;
 RX MEDLINE-99097350; Pubmed-9878253;

RA Petruzzella V., Tiranti V., Fernandez P., Ianna P., Carozzo R.,
 RA Zeviani M.;
 RT "Identification and characterization of human cDNAs specific to BCSL,
 RT PFT12, SCOT, COX15, and COX11, five genes involved in the formation
 RT and function of the mitochondrial respiratory chain.";
 RL Genomics 54:494-504(1998).
 RN [2]
 RP SEQUENCE FROM N.A.

RC TISSUE-BRAIN;
 RX MEDLINE-96207227; Pubmed-8619474;
 RA Anderson B., Wentland M.A., Ricafrente J.Y., Liu W., Gibbs R.A.;
 RT "A 'double adaptor' method for improved shotgun library
 RT construction.";
 RL Anal. Biochem. 236:107-113(1996).
 RN [3]
 RP SEQUENCE FROM N.A.

RC TISSUE-BRAIN;
 RX MEDLINE-97264341; Pubmed-9110174;
 RA Yu W., Anderson B., Worley K.C., Muzny D.M., Ding Y., Liu W.,
 RA Ricafrente J.Y., Wentland M.A., Lennon G., Gibbs R.A.;
 RL "Large-scale concatenation cDNA sequencing.";
 RL Genome Res. 7:353-358(1997).
 DR EMBL: AF026849; AAD08638.1; -
 DR EMBL: AF038195; AAB97365.1; -
 DR INTERPRO: IPR001939; -
 DR PPM: PF00004; AAA; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 419 AA; 47534 MW; 7F0F98BA62F2CBB8 CRC64;

Query Match 42.2%; Score 49; DB 4; Length 419;
 Best Local Similarity 80.0%; Pred. No. 15;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 NPAMYTGRCI 14
 II ||| |||
 DB 211 NPKMYTDRCI 220

RESULT 5
 ID 085605 PRELIMINARY; PRT; 223 AA.
 AC 085605;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE EXCINUCLEASE ABC SUBUNIT A (FRAGMENT).
 GN IVRA.
 OS Prevotella albensis.
 OC Bacteria; CFB group; Bacteroidaceae; Prevotella.
 OX NCBI_TaxID=77768;
 RN [1]
 RP SEQUENCE OF 9-191 FROM N.A.
 RC STRAIN-M384;
 RA Walker N.D., McEwan N.R., Wallace R.J.;
 RL "Prevotella albensis putative uvra gene.";
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF070890; AAC24131.2; -
 DR INTERPRO: IPR001617; -
 DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
 KW ATP-binding; Transport.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 223 AA; 24579 MW; 2563BA0BC2996C9 CRC64;

Query Match 41.4%; Score 48; DB 2; Length 223;
 Best Local Similarity 37.8%; Pred. No. 11;
 Matches 14; Conservative 1; Mismatches 4; Indels 18; Gaps 2;

QY 1 TPDINPMTG-----RCIRPVGR 20
 II ||| ||| |||
 DB 57 TPRSNPATYTGVSFSDIRTLFVGLPEAKIRGYKP-GRR 92

RC STRAIN-A3(2);
 RA Redenbach M., Kleser H.M., Denapalte D., Eichner A., Cullum J.,
 RA Kinsahl H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 DR EMBL: AL21746; CAB57411.1; -
 KW Hypothetical protein.
 SO SEQUENCE 333 AA; 36312 MW; 884FE2861837FD76 CRC64;

Query Match 40.1%; Score 46.5; DB 2; Length 333;
 Best Local Similarity 35.5%; Pred. No. 29;
 Matches 11; Conservative 2; Mismatches 7; Indels 11; Gaps 1;

OY 1 PDINPAMWT-----GRIPIRYGPF 20
 DB 19 TWEEPWRTETAIQWIEARLAHGLPGRW 49

RESULT 10
 O9PH76 PRELIMINARY; PRT; 333 AA.
 AC O9PH76;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE HYDROXYBENZONATE OCTADRENYLTRANSFERASE.
 GN XF0068.
 OS Xylella fastidiosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
 CC Xylella.
 OX NCBI_TaxID=2371;
 RN [1]

SEQUENCE FROM N.A.
 RC STRAIN-9ASC;
 RX MEDLINE=20365717; PubMed=10910347;
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
 RA Alvarenga R., Alves L.M.C., Araya J.E., Baia J.S., Baptista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordoli S., Bove J.M., Briones M.R.S.,
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carter H.,
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 RA Coutinho L.P., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
 RA Faccinini A.P., Ferreira A.J.S., Fereira V.C.A., Ferro J.A.,
 RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
 RA Garner M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 RA Ho P.L., Hohelsel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
 RA Krieger J.E., Kuramae E.E., Laigret F., Lambdas M.R., Leite L.C.C.,
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Martino C.L.,
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Montello-Vitorillo C.B.,
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 RA Nhami A.J.R., Nobrega F.G., Nunes L.R., Oliveira M.A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
 RA Queglio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
 RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuchioka M.H.,
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Zeldanis J., Zerbini J.C.;

RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
 RL Nature 406:151-157(2000).
 DR EMBL: AE003860; AAF82881.1; -
 DR INTERPRO: IPR000537; -
 DR PIRAM: PF01040; UBLA: 1.
 DR PROSITE: PS00943; UBLA: UNKNOWN_1.
 SO SEQUENCE 333 AA; 37931 MW; ECF3F47J6C962B95 CRC64;

Query Match 39.7%; Score 46; DB 2; Length 333;

Best Local Similarity 53.3%; Pred. No. 34;
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 4 INPAMTGRGIRPYG 18
 DB 54 LDPYMKLARDRDPVG 68

RESULT 11

O76383 PRELIMINARY; PRT; 527 AA.

AC O76383;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE C24G6.6 PROTEIN.
 GN C24G6.6.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peleiderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]

SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurtry A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Shownkeen R.,
 RA Smailton N., Smith A., Sonhammer E., Staden R., Sultison J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wollman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RL elegans.";
 RN Nature 368:32-38(1994).

SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 RL EMBL: AF067936; AAC19213.1; -
 DR INTERPRO: IPR000205; -
 DR INTERPRO: IPR002937; -
 DR PIRAM: PF01593; Amino-oxidase; 1.
 SO SEQUENCE 527 AA; 59805 MW; 9FBE1EB84437C5CB CRC64;

Query Match 39.7%; Score 46; DB 5; Length 527;
 Best Local Similarity 58.3%; Pred. No. 56;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 PDINPAMWTGRG 13
 DB 370 PNVLSAWYAGRG 381

RESULT 12

O00050 PRELIMINARY; PRT; 555 AA.

AC O00050;
 DT 01-JUL-1997 (TREMBLrel. 04, Created)
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
 DE TRANSPOSASE.

Query Match 39.7%; Score 46; DB 5; Length 527;
 Best Local Similarity 58.3%; Pred. No. 56;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

RESULT	13			
Q9NE93				
ID	Q9NE93	PRELIMINARY;	PRT;	820 AA.
AC	Q9NE93;			
DT	01-OCT-2000 (TREMBLrel, 15, Created)			
DT	01-OCT-2000 (TREMBLrel, 15, last sequence update)			
DT	01-OCT-2000 (TREMBLrel, 15, last annotation update)			
DE	HYDROTHERMAL 86.4 KDA PROTEIN.			
GN	L787.05.			
OS	Leishmania major.			
OC	Eukaryotes; Eukaryotes; Kinetoplastida; Trypanosomatidae; Leishmania			
OX	NCBI_TaxID=5664;			
RI	[1]			
RN	SEQUENCE FROM N.A.			
RP	STRAIN-FRIEDLIN;			
RC	STRAIN-FRIEDLIN;			
RA	Robben J., Aert R., Volckaert G., Ivens A.C., Quail M.,			
RA	Rajadream M.A., Barrell B.G.;			
RL	Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.			
RL	[2]			
RN	SEQUENCE FROM N.A.			
RP	STRAIN-FRIEDLIN;			
RC	STRAIN-FRIEDLIN;			
RA	Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,			
RA	Smith D.F.;			
RT	"A physical map of the Leishmania major Friedlin genome.";			
RL	Genome Res. 8:135-145(1998).			
RL	EMBL/ AL163492; CAB86689.1; -.			
DR	Hypothetical protein.			
Q5	SEQUENCE 820 AA; 86386 MW; F4C80D9FDE666817 CRC64;			

```

Query Match Similarity      39.7%; Score 46; DB 5; Length 820;
Best Local Similarity      47.1%; Pred No. 92;
Matches      8; Conservative      4; Mismatches      5; Indels      0; Gaps      0;

Qy      2 PDINPAMYTGGRIPVG 18
      1 :|||:|:|:|:|
Db      406 PFTSPALYSGNGSOPLG 422

RESULT 14

```

ID	O9V3E5		PRT:	267 AA.
AC	O9V3E5		PRELIMINARY;	
DT	01-MAY-2000	(TREMBLrel. 13, Created)		
DT	01-MAY-2000	(TREMBLrel. 13, Last sequence update)		
DT	01-JUN-2000	(TREMBLrel. 14, Last annotation update)		
DE	C64999	PROTEIN.		
GN	C64999			
OS	Drosophila melanogaster ('Fruit fly').			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
CC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
CC	Ephyridioidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxId=7227;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BKERELEY;			
RX	MEDLINE=20196006; PubMed=10731132;			
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,			
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,			
RA	Sutton G.C., Wortman J.R., Yeandlel M.D., Zhang Q., Chen L.X.,			
RA	Brandon R.C., Rogers I.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,			
RA	Wan K.H., Doyle C., Baxter E.G., Helt J.G., Nelson C.R., Miklos G.L.G.,			
RA	Abdill J.M., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,			
RA	Bailly R.F., Basu A.I., Bakendate J., Bayraktaroglu I., Beasley E.M.,			
RA	Besson K.Y., Benos P.V., Bernan B.P., Bhargava D., Bolshakov S.,			
RA	Borkova D., Botchan M.R., Bouck T., Broksstein P., Brotlier P.,			
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,			
RA	Cherry J.N., Cavley S., Dahike C., Davenport L.B., Davies P.,			
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,			
RA	Dodson K.C., Dopp L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,			
RA	Dudkin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,			
RA	Foster C., Gabrielian A.E., Garey J.H., Gu Z., Guan P., Hariss M.,			
RA	Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,			
RA	Hoslin D., Houston K.A., Howland T.J., Mei M.-H., Ikegawa C.,			
RA	Jatalai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,			
RA	Kimmel B.E., Kodira C.D., Kratz C., Kravitz S., Kulp D., Lai Z.,			
RA	Lasko P., Lei Y., Lavitsky A.A., Li J., Li Z., Liang Y., Lin X.,			
RA	Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,			
RA	Mekullov G., Milblina N.V., Mobarry C., Morris J., Mosherfi A.,			
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,			
RA	Nelson D.R., Nelson K.A., Nixon K., Nusseren D.R., Paciel J.M.,			
RA	PalazzoLo M., Pitman G.S., Pan S., Pollard J., Puti V., Reese M.G.,			
RA	Reinert K., Remington K., Saunders R.D.C., Schneider F., Shen H.,			
RA	Shue B.C., Stiden-Klamas I., Simpson M., Skipski M.P., Smith T.,			
RA	Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,			
RA	Svylske R., Tecor C., Turner R., Ventler E., Wang A.H., Wang X.,			
RA	Wang Z.-Y., Massaraghi D.A., Welnscock G.M., Weissbach J.,			
RA	Williams S., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,			
RA	Xe J., Yeh R.-F., Zaveri J.S., Zhao M., Zhang G., Zhao G., Zheng L.,			
RA	Zheng X.H., Zhong F.N., Zhuo W., Zhou X., Zhu S., Zhu X., Smith H.O.,			
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;			
RT	"The genome sequence of Drosophila melanogaster.";			
RL	Science 287:2185-2195(2000).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Page-McCaw A.W., Tsang G., Rubin G.M.;			
RT	"Sequencing Drosophila cDNAs related to tetrapansins";			
RL	Submitted (DEC-1999) to the EMBL/Genbank/DDJB databases.			
DR	EMBL; AE003553; AF50354.1;			
DR	EMBL; AF220044; AAFT3828.1;			
DR	LIBRARY; FBgn0035936; CG4999.			
DR	INTERPRO; IPRO00301;			
DR	PFAM; PF00335; transmembrane4; 1.			
DR	PRINTS; PR00259; TMFOUR			
SQ	SEQUENCE 267 AA; 29879 MW; DEB1AB2A59BB65E CRC64;			

Query Match	38.8%	Score 45;	DB 5;	Length 267;
Best Local Similarity	46.7%	Pred. NO. 38;		
Matches	7;	Conservative	3;	Mismatches 5;
				Indels 0;
				Gaps 0;

OY 3 DNPANYTGGIRPV 17
 : : : : :
 Db 171 DASPANYNGKGNRTI 185

RESULT 15

O9NRB7 PRELIMINARY; PRT; 330 AA.
 AC O9NRB7;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE ADAPTER PROTEIN GRID.
 GN GRID.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE-20281669; PubMed-10820259;
 RA Ellis J.H., Ashman C., Burden M.N., Kilpatrick K.E., Morse M.A.,
 RA Hamblin P.A.;
 RT "GRID: a novel Grb-2-related adapter protein that interacts with the
 RT activated T cell costimulatory receptor CD28.";
 RL J. Immunol. 164:5805-5814(2000).
 DR EMBL; AF236119; AAF60319.1; "-"
 SQ SEQUENCE 330 AA; 37943 MW; 74F4C8D849B56D55 CRC64;

Query Match

Best Local Similarity 38.8%; Score 45; DB 4; Length 330;

Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 5 NPAWYTGGRPVGRF 20
 ||:|:|:|:|:|:|

Db 305 NPSWMTGRILNKLGF 320

Search completed: April 17, 2001, 15:48:08
 Job time: 561 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 17, 2001, 15:45:54 ; Search time 70.08 seconds

(Without alignments)
19.613 Million cell updates/sec

Title: US-09-446-543a-50
Perfect score: 116
Sequence: 1 TP0INPAWTGRCIRPVGRF 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR_67:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
5: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	51	44.0	962	2	H69157	excinnuclease ABC c
2	49	42.2	1670	2	S71551	DNA-directed DNA p
3	48.5	41.8	664	2	F83376	conserved hypotet
4	48	41.4	790	2	T47959	hypothetical prote
5	47	40.5	430	1	B69009	conserved hypotet
6	47	40.5	453	2	S18597	tubulin beta chain
7	47	40.5	719	2	S61046	ABPI protein - Yea
8	47	40.5	940	2	A82329	excinnuclease ABC,
9	47	40.5	972	2	A70619	excinnuclease ABC c
10	46.5	40.1	652	1	S41522	tetracycline resis
11	46	39.7	330	2	JE0376	Grb-2 related adap
12	46	39.7	333	2	H82852	hydroxybenzoate oc
13	46	39.7	527	2	T33175	hypothetical prote
14	46	39.7	943	2	D64057	excinnuclease ABC c
15	46	38.8	482	1	S40887	RNS167 protein - Y
16	45	38.8	767	2	T21969	hypothetical prote
17	45	38.8	798	2	S11210	probable untr prote
18	45	38.8	940	1	BVECUA	excinnuclease ABC c
19	45	38.8	953	2	D71645	excinnuclease ABC c
20	45	38.8	965	2	C82560	excinnuclease ABC s
21	44	37.9	105	2	D81319	hypothetical prote
22	44	37.9	184	2	E72248	ribosomal protein
23	44	37.9	418	2	F75587	probable glycosylt
24	44	37.9	548	2	T47548	hypothetical prote
25	44	37.9	771	2	C75023	DNA polymerase I p
26	44	37.9	945	2	B83117	excinnuclease ABC s
27	44	37.9	952	1	T46550	excinnuclease ABC c
28	44	37.9	970	2	S77349	excinnuclease ABC c
29	44	37.9	1014	2	T36031	excinnuclease ABC c

30	44	37.9	1312	2	S68593	DNA-directed DNA p
31	43.5	37.5	506	2	H83396	probable aldehyde
32	43.5	37.5	506	2	F83142	probable aldehyde
33	43.5	37.5	1501	2	T45623	hypothetical prote
34	43.5	37.5	1680	2	A43434	furin (EC 3.4.21.7
35	43	37.1	220	2	C83292	probable glucatino
36	43	37.1	241	1	S64445	hypothetical prote
37	43	37.1	309	2	T32376	hypothetical prote
38	43	37.1	476	2	G64720	probable amino aci
39	43	37.1	498	2	T05021	beta-glucosidase h
40	43	37.1	503	2	A82193	Sun/nucleolar prot
41	43	37.1	517	2	T05022	beta-glucosidase h
42	43	37.1	948	2	B81883	excinnuclease ABC c
43	43	37.1	949	2	A81138	excinnuclease ABC c
44	43	37.1	960	2	A71315	excinnuclease ABC c
45	43	37.1	1282	2	T30577	DNA topoisomerase

ALIGNMENTS

RESULT 1
H69157
excinnuclease ABC chain A - Methanobacterium thermoautotrophicum (strain Delta N)
N:Alternate names: uvra protein
C:Contains: excision endonuclease ABC (EC 3.1.-.-) chain A
C:Species: Methanobacterium thermoautotrophicum
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 12-Nov-1999
C:Accession: H69157
R:Smith, D.R.; Doucette-Stamm, L.A.; Delonghery, C.; Lee, H.; Dubois, J.; Aldredge, T.
; Qiu, D.; Spadafora, R.; Vicalire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani,
K.I.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A>Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fu
A:Reference number: A69000; MID:98037514
A:Accession: H69157
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-962 <MTB>
A:Cross-references: GB:AE000828; GB:AE000666; NID:g2621504; PIDN:AAB84949.1; PID:g262
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTN443
A:Start codon: TTG
C:Superfamily: excinnuclease ABC chain A; ATP-binding cassette homology
C:Keywords: ATP; DNA binding; DNA repair; duplication; hydrolase; P-loop
F:38-45/Region: nucleotide-binding motif A (P-loop)
F:632-915/Domain: ATP-binding cassette homology (ABCE)
F:649-656/Region: nucleotide-binding motif A (P-loop)

Query Match 44.0%; Score 51; DB 2; Length 962;
Best Local Similarity 40.5%; Pred. No. 9.8;
Matches 15; Conservative 0; Mismatches 4; Indels 18; Gaps 2;

QY 1 TP0INPAWTG-----RGIRPVGRF 20
Db 704 TP0INPAWTGVTTHRELPQPEARKRGVRF 729

RESULT 2
S71551
DNA-directed DNA polymerase (EC 2.7.7.7) KOD, intein containing precursor - Pyrococcus
C:Species: Pyrococcus sp.
A:Variety: strain KOD1
C>Date: 22-Nov-1996 #sequence_revision 22-Nov-1996 #text_change 20-Apr-2000
C:Accession: S71551
R:Kakihara, H.; Takagi, M.; Imanaka, T.
submitted to the EMBL Data Library, March 1994
A:Description: Cloning and analysis of the DNA polymerase gene from a new hyperthermo
A:Reference number: S71551
A:Accession: S71551

Best Local Similarity 70.0%; Pred. No. 18;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 8 WYTGRCIRPV 17
IIII I:II
Db 399 WYTGRCIRPV 408

RESULT 7

ABP1 protein - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein D1478; protein YDL167c
C:Species: Saccharomyces cerevisiae
C:Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 29-Oct-1999
C:Accession: S61046; S31139; S67719
R:Poln, T.M.
submitted to the EMBL Data Library, November 1995
A:Reference number: S61010
A:Accession: S61046
A:Molecule type: DNA
A:Residues: 1-719 <POS>
A:Cross-references: EMBL:Z67750; NID:g1061256; PTDN:CAA91579.1; PID:g1061272
R:Wehner, E.P.; Rao, E.; Brendel, M.
Mol. Gen. Genet. 237, 351-358, 1993
A:Title: Molecular structure and genetic regulation of SFA, a gene responsible for resis
A:Reference number: S31138; MUID:93247548
A:Accession: S31139
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-492, N', 494-719 <NEG>
A:Cross-references: EMBL:X68020; NID:9577609; PTDN:CAA48159.1; PID:9288590
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1992
R:Poln, T.M.
submitted to the Protein Sequence Database, July 1996
A:Reference number: S67708
A:Accession: S67719
A:Molecule type: DNA
A:Residues: 1-719 <POS>
A:Cross-references: EMBL:Z74215; NID:g1431265; PTDN:CAA98741.1; PID:e253076; PID:g143126
A:Experimental source: strain S288C
C:Genetics:
A:Gene: SGD:NRP1; ARP1
A:Cross-references: MIPS:YDL167c; SGD:S0002326
A:Map position: 4L

Query Match 40.5%; Score 47; DB 2; Length 719;
Best Local Similarity 58.3%; Pred. No. 29;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 7 AWYTGRCIRPV 18
.:.:I:IIIIII
Db 244 SWYTGRCIRPV 255

RESULT 8

excinuclease ABC, chain A VC0394 [imported] - Vibrio cholerae (group O1 strain N16961)
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 15-Sep-2000
C:Accession: A82329
R:Heideberg, J.F.; Eissen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, B.; Dragol, I.; Sellers, F.
L. R.R.; Metcalne, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833
A:Accession: A82329
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-940 <HEI>
A:Cross-references: GB:AE004127; GB:AE003852; NID:9654808; PTDN:AAF93567.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:
A:Gene: VC0394
A:Map Position: 1
C:Superfamily: excinuclease ABC chain A: ATP-binding cassette homology

Query Match 40.5%; Score 47; DB 2; Length 940;
Best Local Similarity 37.8%; Pred. No. 39;
Matches 14; Conservative 1; Mismatches 4; Indels 18; Gaps 2;

QY 1 TPDPNPATYTG-----RGIRPVGRF 20
II IIIIIII
Db 695 TPSPNPATYTGIFTPRELPAGTQESSRKGYP-GRF 730

RESULT 9

A70619
excinuclease ABC chain A - Mycobacterium tuberculosis (strain H37RV)
N:Alternate names: uvra protein
M:Contains: excision endonuclease ABC (EC 3.1.-.-) chain A
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: A70619
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
Connor, R.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrett, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A:Reference number: A70500; MUID:98295987
A:Accession: A70619
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-972 <COL>
A:Cross-references: GB:Z85982; GB:A1123456; NID:93261718; PTDN:CA80653.1; PID:g18389
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: uvra
C:Superfamily: excinuclease ABC chain A: ATP-binding cassette homology
C:Keywords: ATP; DNA binding; DNA repair; hydrolase
F:32-39/Region: nucleotide-binding motif A (P-loop)
F:637-920/Domain: ATP-binding cassette homology <ABCE>
F:654-661/Region: nucleotide-binding motif A (P-loop)

Query Match 40.5%; Score 47; DB 2; Length 972;
Best Local Similarity 37.8%; Pred. No. 40;
Matches 14; Conservative 1; Mismatches 4; Indels 18; Gaps 2;
QY 1 TPDPNPATYTG-----RGIRPVGRF 20
II IIIIIII
Db 709 TPSPNPATYTGIFTPRELPAGTQESSRKGYP-GRF 744

RESULT 10

S41522
tetracycline resistance protein tetB - Clostridium perfringens
C:Species: Clostridium perfringens
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S41522
R:Sloan, J.; McMurtry, L.M.; Lyrae, D.; Levy, S.B.; Rood, J.I.
Mol. Microbiol. 11, 403-415, 1994
A:Title: The Clostridium perfringens Tet P determinant comprises two overlapping gene
family of tetracycline-resistance determinants.
A:Reference number: S41522; MUID:94224158
A:Accession: S41522
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-652 <SLC>
A:Cross-references: GB:U20800; NID:9456030; PTDN:AAA20117.1; PID:g456035
C:Superfamily: translation elongation factor G; translation elongation factor Tu homo
C:Keywords: antibiotic resistance; GTP binding; P-loop
F:5-132/Domain: translation elongation factor Tu homology <ETU>

C:Genetics:
 A:Gene: uvra
 C:Function:
 A:Description: has ATPase and DNA binding activity; involved in DNA repair
 C:Superfamily: excinuclease ABC chain A; ATP-binding cassette homology
 C:Keywords: ATP; DNA binding; DNA repair; duplication; hydrolase; P-loop
 F:3138/Region: nucleotide-binding motif A (P-loop)
 F:623-907/Domain: ATP-binding cassette homology <ABCE>
 F:640-647/Region: nucleotide-binding motif A (P-loop)

Query Match 39.7%; Score 46; DB 2; Length 943;
 Best Local Similarity 37.6%; Pred. No. 55;
 Matches 14; Conservative 0; Mismatches 5; Indels 18; Gaps 2;

QY 1 TPDINPAMWTG-----RGIRPYGRF 20
 |||||
 Db 696 TPRSNPATYTGTFPIRELFGVPEARAKGYNP-GRF 731

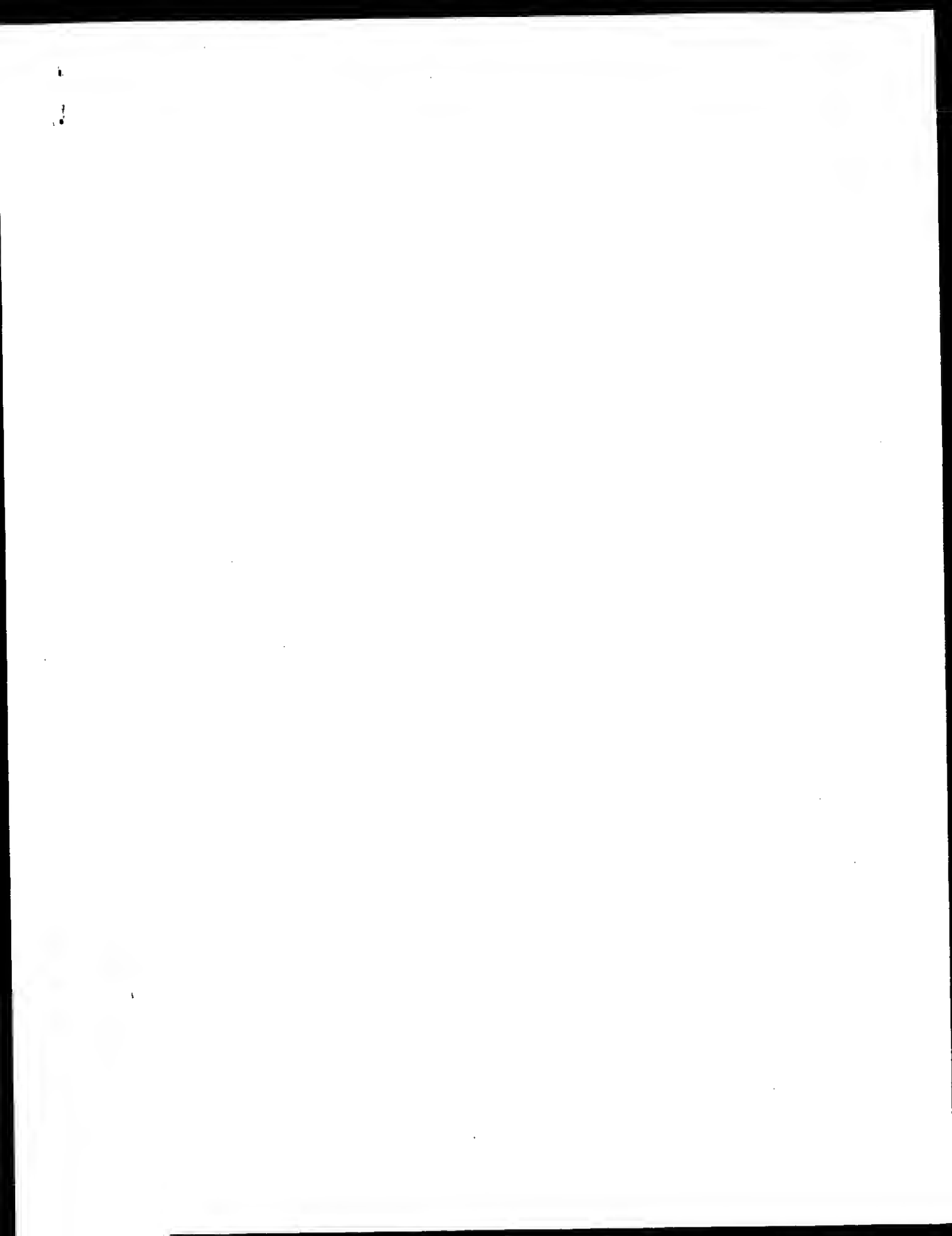
RESULT 15

S40887
 RVSI67 protein - yeast (Saccharomyces cerevisiae)
 N:Alternate names: protein YDR388w
 C:Species: Saccharomyces cerevisiae
 C>Date: 31-Mar-1992 #sequence_revision 06-Feb-1995 #text_change 21-Jul-2000
 C:Accession: S40887; S69672
 R:Baer, F.; Urdaci, M.; Agile, M.; Crouzet, M.
 Mol. Cell. Biol. 13, 5070-5084, 1993
 A>Title: Alteration of a yeast SH3 protein leads to conditional viability with defects in
 A:Reference number: S40887; MUID:93330293
 A:Accession: S40887
 A:Molecule type: DNA
 A:Residues: 1-482 <BAU>
 A:Cross-references: EMBL:M92092; NID:q172615; PIDN:AAA35051.1; PID:q172616
 R:Dietrich, F.S.
 Submitted to the EMBL Data Library, July 1995
 A:Description: The sequence of S. cerevisiae cosmid 9481, 9509, 9926, 9461, and lambda
 A:Reference number: S69665
 A:Accession: S69672
 A:Molecule type: DNA
 A:Residues: 1-482 <DIE>
 A:Cross-references: EMBL:U32274; NID:g927313; PIDN:AAB64830.1; PID:g927321; GSPDB:GN0000
 C:Genetics:
 A:Gene: SGD:RVSI67; MIPS:YDR388w
 A:Cross-references: SGD:S0002796; MIPS:YDR388w
 A:Map position: 4R
 C:Superfamily: RVSI67 protein; RVSI61 protein homology; SH3 homology
 C:Keywords: Transmembrane protein
 F:4-270/Domain: RVSI61 protein homology <RVSI>
 F:292-422/Region: alanine/glycine/proline-rich
 F:426-477/Domain: SH3 homology <SH3>

Query Match 38.8%; Score 45; DB 1; Length 482;
 Best Local Similarity 47.6%; Pred. No. 39;
 Matches 10; Conservative 3; Mismatches 2; Indels 6; Gaps 2;

QY 1 TPDINPAMWTG-----GIRP 16
 |||||
 Db 454 TPDVN-EMWTGKNGCGGVFP 473

Search completed: April 17, 2001, 15:45:55
 Job time: 603 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 17, 2001, 15:38:42 ; Search time 116.94 Seconds
(without alignments)
9.777 Million cell updates/sec

Title: US-09-446-543a-50
Perfect score: 116
Sequence: 1 FPDINPAMWTGGRIRPVGRF 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: /SIDSL1/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDSL1/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDSL1/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDSL1/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDSL1/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDSL1/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDSL1/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDSL1/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDSL1/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDSL1/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDSL1/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDSL1/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDSL1/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDSL1/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDSL1/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDSL1/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDSL1/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDSL1/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDSL1/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDSL1/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDSL1/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDSL1/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	116	100.0	20	18 W31387	Rat type G protein
2	116	100.0	20	20 W97234	Rat type ligand po
3	116	100.0	20	20 W95175	Murine pituitary-d
4	116	100.0	20	21 B10358	Rat oxytocin secre
5	116	100.0	20	21 Y49302	19P2 ligand peptid
6	116	100.0	21	18 W31388	Rat type G protein
7	116	100.0	21	21 B10359	Rat oxytocin secre
8	116	100.0	22	18 W31389	Rat type G protein
9	116	100.0	22	21 B10360	Rat oxytocin secre
10	116	100.0	31	18 W31384	Rat type G protein
11	116	100.0	31	20 W97233	Rat type ligand po

12	116	100.0	31	20 W87614	Rat 19P2 ligand.
13	116	100.0	31	20 W95173	Murine pituitary-d
14	116	100.0	31	20 W95174	Murine pituitary-d
15	116	100.0	31	21 B10355	Rat oxytocin secre
16	116	100.0	31	21 Y87504	Rat prolactin-rele
17	116	100.0	31	21 Y49292	19P2 ligand peptid
18	116	100.0	32	18 W31385	Rat type G protein
19	116	100.0	32	21 B10356	Rat oxytocin secre
20	116	100.0	33	18 W31386	Rat type G protein
21	116	100.0	33	21 B10357	Rat oxytocin secre
22	116	100.0	82	20 W95172	Murine pituitary-d
23	116	100.0	83	18 W31383	Rat type G protein
24	116	100.0	83	20 W97225	Rat type ligand po
25	116	100.0	83	21 B10354	Rat oxytocin secre
26	111	95.7	20	18 W31374	Bovine G protein-c
27	111	95.7	20	20 W97232	Bovine pituitary-d
28	111	95.7	20	20 W95191	Bovine pituitary-d
29	111	95.7	20	21 B10350	Bovine oxytocin se
30	111	95.7	20	21 Y49301	Bovine G protein-c
31	111	95.7	21	18 W31375	Bovine G protein-c
32	111	95.7	21	20 W95192	Bovine pituitary-d
33	111	95.7	21	21 B10351	Bovine oxytocin se
34	111	95.7	22	18 W31376	Bovine G protein-c
35	111	95.7	22	20 W95193	Bovine pituitary-d
36	111	95.7	22	20 B10352	Bovine oxytocin se
37	111	95.7	31	18 W31371	Bovine G protein-c
38	111	95.7	31	20 W97218	Bovine pituitary-d
39	111	95.7	31	20 W87613	Bovine 19P2 ligand
40	111	95.7	31	20 W95188	Bovine pituitary-d
41	111	95.7	31	21 B10347	Bovine oxytocin se
42	111	95.7	31	21 Y49290	19P2 ligand peptid
43	111	95.7	31	21 Y49298	19P2 ligand peptid
44	111	95.7	32	18 W31372	Bovine G protein-c
45	111	95.7	32	20 W95189	Bovine pituitary-d

ALIGNMENTS

RESULT 1	
W31387	standard; Peptide: 20 AA.
XX	
AC	W31387;
XX	
DT	06-APR-1998 (first entry)
XX	
DE	Rat type G protein-coupled receptor ligand fragment 4.
XX	
KW	G protein-coupled receptor; ligand binding; Pharmaceutical;
KW	modulator; pituitary; central nervous system; pancreas; prophylactic;
KW	therapeutic agent.
XX	
OS	Rat sp.
XX	
PN	W09724436-A2.
XX	
PD	10-JUL-1997.
XX	
PF	26-DEC-1996; 96MO-JP03821.
XX	
PR	18-SEP-1996; 96JP-0246573.
PR	28-DEC-1995; 95JP-034371.
PR	15-MAR-1996; 96JP-0059419.
PR	12-AUG-1996; 96JP-0211805.
XX	
PA	(TAKE) TAKEDA CHEM IND LTD.
XX	
PI	Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;
PI	Kawanata Y, Kitada C;
XX	
DR	WPI; 1997-363672/33.
DR	N-PDB; V02424.

XX L19and peptide for G protein-coupled receptor - acts by modulating
 PT function in the central nervous system, pancreas and pituitary gland
 XX
 PS Claim 2; Page 180; 258pp; English.
 CC This sequence represents a peptide fragment from a novel rat type
 CC ligand polypeptide corresponding to amino acid residues 33 to 52 of the
 CC sequence represented in W91383 and is used in an assay to monitor ligand
 CC binding to the G protein-coupled receptor protein. Pharmaceutical
 CC compositions containing this ligand may be used as a pituitary function
 CC modulator, a central nervous system modulator or a pancreatic function
 CC modulator. This ligand could have specific applications as a prophylactic
 CC or therapeutic agent for dementia, depression, hyperkinetic syndrome,
 CC disturbance of consciousness, anxiety syndrome, schizophrenia, trauma,
 CC growth hormone secretory disease, hyper- and polyphagia, hyperlipidaemia,
 CC hypercholesterolaemia, hyperglycaemia, hyperprolactinaemia, diabetes,
 CC cancer, pancreatitis, renal disease, Turner's syndrome, neurosis, asthma,
 CC rheumatoid arthritis, spinal injury, transient brain ischaemia, epilepsy,
 CC amyotrophic lateral sclerosis, acute myocardial infarction, infertility,
 CC spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis,
 CC osteoporosis and/or oligogalactia. Assays can also be developed to screen
 CC compounds which are capable of altering the binding activity of the
 CC ligand affecting activation of the G protein-coupled receptor protein.
 XX
 SO Sequence 20 AA;
 Query Match 100.0%; Score 116; DB 18; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.3e-11;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 TPDINPAMWTGRCIRPVGRF 20
 Db 1 tpdinpwtytgrtprvgrf 20
 RESULT 2
 W97234
 ID W97234 standard; peptide; 20 AA.
 AC W97234;
 DT 06-MAY-1999 (first entry).
 DE Rat type ligand polypeptide fragment.
 XX
 XX Rat type ligand; modulation; prolactin secretion;
 KW G protein-coupled receptor; GPCR; hypocoovarianism; gonocyst cacogenesis;
 KW menopausal syndrome; euthyroid; hypometabolism; lactation;
 KW pituitary adenomatosis; brain tumour; emmenopathy; autoimmune disease;
 KW prolactinoma; infertility; impotence; amenorrhea; galactorrhea;
 KW acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;
 KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dysospermia;
 KW contraceptive; placental function; choriocarcinoma; hydatid mole;
 KW interruption mole; abortion; unfertility fetus; abnormal saccharometabolism;
 KW abnormal lipidmetabolism; oxytocia.
 XX
 OS Rattus sp.
 PN MO9858962-A1.
 PD 30-DEC-1998.
 XX
 XX 22-JUN-1998; 98WO-JP02765.
 XX 23-JUN-1997; 97JP-0165437.
 XX (TAKE) TAKEDA CHEM IND LTD.
 PA Fujii R, Hinuma S, Kawamata Y, Matsumoto H;
 PI WPI; 1999-105614/09.
 DR

XX Use of G protein-coupled receptor ligands - for modulating prolactin
 PT secretion or placental function, e.g. for treating menopausal
 PT syndrome, tumours, autoimmune disease or abnormal pregnancy
 XX
 PS Claim 3; Page 154; 241pp; English.
 CC The present sequence represents a rat type ligand fragment. It
 CC is used in the course of the invention. The specification describes
 CC an agent for modulating prolactin secretion which comprises a
 CC ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)
 CC protein. The agents for promoting prolactin secretion can be used for
 CC treating or preventing hypocoovarianism, gonocyst cacogenesis, menopausal
 CC syndrome, euthyroid or hypometabolism. They can be used for promoting
 CC lactation in a domestic mammal and as an aphrodisiac. The agents for
 CC inhibiting prolactin secretion can be used for treating or preventing
 CC pituitary adenomatosis, brain tumour, emmenopathy, autoimmune disease,
 CC prolactinoma, infertility, impotence, amenorrhea, galactorrhea,
 CC acromegaly, Chiari-Frommel syndrome, Argonz-del Castillo syndrome,
 CC Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dysospermia.
 CC The inhibitory agents can also be used as contraceptives. The agents for
 CC modulating placental function can be used for treating or preventing
 CC choriocarcinoma, hydatid mole, interruption mole, abortion, unfertility fetus,
 CC abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.
 XX
 SO Sequence 20 AA;
 Query Match 100.0%; Score 116; DB 20; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.3e-11;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 TPDINPAMWTGRCIRPVGRF 20
 Db 1 tpdinpwtytgrtprvgrf 20
 RESULT 3
 W95175
 ID W95175 standard; Protein; 20 AA.
 AC W95175;
 DT 10-MAR-1999 (first entry)
 DE Murine pituitary-derived ligand polypeptide antigenic epitope.
 XX
 XX pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;
 KW GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas;
 KW tissue; screen; therapeutic; binding; senile dementia; ligand; murine;
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;
 KW Creutzfeld-Jakob disease; poisoning; schizophrenia; growth hormone;
 KW secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;
 KW gene therapy; transgenic animal; epitope.
 XX
 OS Mus sp.
 PN MO9849295-A1.
 PD 05-NOV-1998.
 XX
 XX 27-APR-1998; 98WO-JP01923.
 XX 28-APR-1997; 97JP-0109974.
 XX (TAKE) TAKEDA CHEM IND LTD.
 PA Fukusumi S, Hinuma S;
 PI WPI; 1999-009423/01.
 DR
 XX New polypeptide ligand for orphan G protein coupled receptors - used
 PT for treating disorders of central nervous system, pituitary and

PT pancreas, and for drug screening
 XX
 PS Disclosure; Page 26; 206pp; English.
 XX
 CC The invention relates to a murine pituitary-derived ligand polypeptide
 CC which is a ligand for the G-protein coupled orphan receptor designated
 CC GPR10 (human) or OHR-1 (rat). Cells transformed with a vector containing
 CC the ligand polypeptide encoding DNA are used to produce a recombinant
 CC ligand polypeptide. The ligand polypeptide, and its fragments, modulate
 CC function of the pituitary, central nervous system, pancreas and other
 CC tissues and can be used to screen for agents that modulate binding of the
 CC polypeptide to the receptor; to quantify the amount of receptor in a
 CC sample and to raise antibodies. They may also be used therapeutically,
 CC e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's
 CC diseases; Creutzfeldt-Jakob disease; poisoning by heavy metals or drugs;
 CC diabetes; schizophrenia; disorders of growth hormone secretion; cancer;
 CC rheumatoid arthritis, epilepsy and many others, also to improve post-
 CC operative nutritional status and as vasopressor. Transgenic animals
 CC carrying the ligand polypeptide encoding DNA or its mutant are used to
 CC study the function of the polypeptide-expressing genes, as models of
 CC disease, for drug screening and as source of cell lines. The ligand
 CC polypeptide DNA is used as a source of probes and primers; to identify
 CC related sequences; in receptor-binding assays; for production of Ab and
 CC antisera; in drug development; for gene therapy and to develop transgenic
 CC animals. Sequences W95174 to W95178 represent antigenic epitopes which
 CC can be used for the preparation of anti-ligand polypeptide antibody.
 XX
 SQ Sequence 20 AA;
 Query Match 100.0%; Score 116; DB 20; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.3e-11;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TPDINPAMYTGKIRPVGRF 20
 |||
 Db 1 tpdinpawytgtrgtrpygrf 20
 RESULT 4
 B10358
 ID B10358 standard; peptide; 20 AA.
 AC B10358;
 XX
 DT 24-NOV-2000 (first entry)
 XX
 DE Rat oxytocin secretion promoting peptide SEQ ID NO: 21.
 XX
 KW Rat; oxytocin secretion promoter; G protein-coupled receptor protein;
 KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;
 KW caesarean section; artificial fertilization; galactostasis; goat; pig;
 KW veterinary medicine; milk production.
 XX
 OS Rattus sp.
 XX
 PN MO200038704-A1.
 PD 06-JUL-2000.
 XX
 PF 22-DEC-1999; 99WO-JP07199.
 XX
 PR 25-DEC-1998; 98JP-0369585.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX
 PI Matsumoto H, Kitada C, Hinuma S;
 XX
 DR WPI: 2000-452298/39.
 PT Physiologically-active polypeptide recognized as ligand by G
 PT protein-coupled receptor protein, for promoting secretion of oxytocin,
 PT as drugs for diseases relating to oxytocin secretion and in veterinary

PT medicine
 XX
 PS Claim 5; Page 58; 72pp; Japanese.
 XX
 CC This invention describes a novel oxytocin secretion-regulating agent
 CC which contains a ligand peptide or its salt for the G protein-coupled
 CC receptor protein. It is useful in the form of drugs for ameliorating,
 CC preventing and treating diseases relating to oxytocin secretion e.g.,
 CC weak pains and atonic bleeding, before and after expulsion of placenta,
 CC uterine recovery failure, caesarean section, stoppage of artificial
 CC fertilization or galactostasis and is also applicable in veterinary
 CC medicine for promoting milk production in cow, goat and pig. This
 CC sequence represents a rat peptide which acts as an oxytocin secretion
 CC promoter.
 XX
 SQ Sequence 20 AA;
 Query Match 100.0%; Score 116; DB 21; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.3e-11;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TPDINPAMYTGKIRPVGRF 20
 |||
 Db 1 tpdinpawytgtrgtrpygrf 20
 RESULT 5
 Y49302
 ID Y49302 standard; peptide; 20 AA.
 XX
 AC Y49302;
 XX
 DT 22-FEB-2000 (first entry)
 XX
 DE 19p2 ligand peptide fragment.
 XX
 KW Monoclonal antibody; 19p2 ligand; diagnosis; prolactin secretion;
 KW pituitary; regulatory mechanism; central nervous system; pancreatic.
 XX
 OS Rattus sp.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 20
 FT /note="C-terminal amide"
 XX
 PN MO9960112-A1.
 XX
 PD 25-NOV-1999.
 XX
 PF 20-MAY-1999; 99WO-JP02650.
 XX
 PR 21-MAY-1998; 98JP-0140293.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX
 PI Matsumoto H, Kitada C, Hinuma S;
 XX
 DR WPI: 2000-039381/03.
 XX
 PT New monoclonal antibodies, useful in diagnosis, as drugs and in
 PT studying diseases related to ligand abnormality
 XX
 PS Disclosure; Page 27; 73pp; Japanese.
 XX
 CC The invention provides a monoclonal antibody which has a specific
 CC reaction with the part peptide of the C-terminal of 19p2 ligand or its
 CC derivative. The antibodies can be used in diagnosis or to treat or
 CC prevent diseases associated with abnormality in the pituitary function
 CC regulatory mechanism (e.g. promotion of prolactin secretion), central
 CC nervous regulatory mechanism, and pancreatic function regulatory
 CC mechanism. The antibody-based immunoassay can also be applied in
 CC clarifying the physiological functions of the ligand and its derivative.

CC Sequences Y49290-302 represent peptide fragments of the 1992 ligand.
 XX
 SQ Sequence 20 AA:

Query Match 100.0%; Score 116; DB 21; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.3e-11;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TPDINPAMYTGRCIRPVGRF 20
 |||
 Db 1 tpdinpaawytrgrirpvgrf 20

RESULT 6
 W31388
 ID W31388 standard; peptide: 21 AA.
 AC W31388;
 XX
 XX

DT 06-APR-1998 (first entry)
 XX

DE Rat type G protein-coupled receptor ligand fragment 5.

KM G protein-coupled receptor; ligand binding; pharmaceutical;
 KW modulator; pituitary; central nervous system; pancreas; prophylactic;
 KW therapeutic agent.
 XX
 XX

OS Rat sp.
 XX

PN W09724436-A2.
 XX

PD 10-JUL-1997.
 XX

PE 26-DEC-1996; 96WO-JP03821.
 XX

PR 18-SEP-1996; 96JP-0246573.
 XX

PR 28-DEC-1995; 95JP-034371.
 XX

PR 15-MAR-1996; 96JP-0059419.
 XX

PR 12-AUG-1996; 96JP-0211805.
 XX

PA (TAKE) TAKEDA CHEM IND LTD.
 XX

PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya N;
 XX

PI Kawamata Y, Kitada C;
 XX

DR WPI: 1997-365672/33.
 XX

DR N-PSDB: Y02425.
 XX

PS Claim 2; Page 180; 258pp; English.
 XX

CC This sequence represents a peptide fragment from a novel rat type
 CC ligand polypeptide corresponding to amino acid residues 33 to 53 of the
 CC sequence represented in W31383 and is used in an assay to monitor ligand
 CC binding to the G protein-coupled receptor protein. Pharmaceutical
 CC compositions containing this ligand may be used as a pituitary function
 CC modulator, a central nervous system modulator or a pancreatic function
 CC modulator. This ligand could have specific applications as a prophylactic
 CC or therapeutic agent for dementia, anxiety syndrome, schizophrenia, trauma,
 CC disturbance of consciousness, depression, hyperkinesia, schizophrenia,
 CC growth hormone secretory disease, hyper- and polyphagia, hyperlipidaemia,
 CC hypercholesterolaemia, hyperglycaemia, hyperprolactinaemia, diabetes,
 CC cancer, pancreatitis, renal disease, Turner's syndrome, neurosis, asthma,
 CC rheumatoid arthritis, spinal injury, transient brain ischaemia, epilepsy,
 CC amyloidotic lateral sclerosis, acute myocardial infarction, infertility,
 CC spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis,
 CC osteoporosis and/or oligosacchara. Assays can also be developed to screen
 CC compounds which are capable of altering the binding activity of the
 CC ligand affecting activation of the G protein-coupled receptor protein.
 XX

SQ Sequence 21 AA:

Query Match 100.0%; Score 116; DB 18; Length 21;
 Best Local Similarity 100.0%; Pred. No. 1.3e-11;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TPDINPAMYTGRCIRPVGRF 20
 |||
 Db 1 tpdinpaawytrgrirpvgrf 20

RESULT 7
 B10359
 ID B10359 standard; peptide: 21 AA.
 XX
 XX

AC B10359;
 XX

DT 24-NOV-2000 (first entry)
 XX

DE Rat oxytocin secretion promoting peptide SEQ ID NO: 22.

KM Rat, oxytocin secretion promoter; G protein-coupled receptor protein;
 KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;
 KW caesarean section; artificial fertilization; galactostasis; goat; pig;
 KW veterinary medicine; milk production.
 XX
 XX

OS Rattus sp.
 XX

PN W0200038704-A1.
 XX

PD 06-JUL-2000.
 XX

PE 22-DEC-1999; 99WO-JP07199.
 XX

PR 25-DEC-1998; 98JP-0369585.
 XX

PR (TAKE) TAKEDA CHEM IND LTD.
 XX

PI Matsumoto H, Kitada C, Hinuma S;
 XX

PI WPI: 2000-452298/39.
 XX

PT Physiologically-active polypeptide recognized as ligand by G
 XX

PT protein-coupled receptor protein, for promoting secretion of oxytocin.
 XX

PT as drugs for diseases relating to oxytocin secretion and in veterinary
 XX

PT medicine
 XX

PS Disclosure; Page 58; 72pp; Japanese.
 XX

CC This invention describes a novel oxytocin secretion-regulating agent
 CC which contains a ligand peptide or its salt for the G protein-coupled
 CC receptor protein. It is useful in the form of drugs for ameliorating,
 CC preventing and treating diseases relating to oxytocin secretion e.g.
 CC weak pains and atonic bleeding, before and after expulsion of placenta,
 CC uterine recovery failure, caesarean section, stoppage of artificial
 CC fertilization or galactostasis and is also applicable in veterinary
 CC medicine for promoting milk production in cow, goat and pig. This
 CC sequence represents a rat peptide which acts as an oxytocin secretion
 CC promoter.
 XX

CC sequence 21 AA;
 XX

SQ Sequence 21 AA:

Query Match 100.0%; Score 116; DB 21; Length 21;
 Best Local Similarity 100.0%; Pred. No. 1.3e-11;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TPDINPAMYTGRCIRPVGRF 20
 |||
 Db 1 tpdinpaawytrgrirpvgrf 20

RESULT 8
 ID W31389 standard; Peptide: 22 AA.
 AC W31389:
 DT 06-APR-1998 (first entry)
 DE Rat type G protein-coupled receptor ligand fragment 6.
 XX
 KW G protein-coupled receptor; ligand binding; pharmaceutical;
 KW modulator; pituitary; central nervous system; pancreas; prophylactic;
 KW therapeutic agent.
 XX
 OS Rat sp.
 XX
 PN W09724436-A2.
 PD 10-JUL-1997.
 XX
 PF 26-DEC-1996; 96WO-JP03821.
 XX
 PR 18-SEP-1996; 96JP-0246573.
 PR 28-DEC-1995; 95JP-0343371.
 PR 15-MAR-1996; 96JP-0059419.
 PR 12-AUG-1996; 96JP-0211805.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M,
 PI Kawamata Y, Kitada C;
 DR WPI: 1997-363672/33.
 DR N-PSDB; V02426.
 XX
 PT Ligand peptide for G protein-coupled receptor - acts by modulating
 PT function in the central nervous system, pancreas and pituitary gland
 XX
 PS Claim 2: Page 181; 258pp; English.
 XX
 CC This sequence represents a peptide fragment from a novel rat type
 CC ligand polypeptide corresponding to amino acid residues 33 to 54 of the
 CC sequence represented in W31383 and is used in an assay to monitor ligand
 CC binding to the G protein-coupled receptor protein. Pharmaceutical
 CC compositions containing this ligand may be used as a pituitary function
 CC modulator, a central nervous system modulator or a pancreatic function
 CC modulator. This ligand could have specific applications as a prophylactic
 CC or therapeutic agent for dementia, depression, hyperkinetic syndrome,
 CC disturbance of consciousness, anxiety syndrome, schizophrenia, trauma,
 CC growth hormone secretory disease, hyper- and polyphagia, hyperlipidaemia,
 CC hypercholesterolaemia, hyperglycaemia, hyperprolactinaemia, diabetes,
 CC cancer, pancreatitis, renal disease, Turner's syndrome, neurosis, asthma,
 CC rheumatoid arthritis, spinal injury, transient brain ischaemia, epilepsy,
 CC amyotrophic lateral sclerosis, acute myocardial infarction, infertility,
 CC spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis,
 CC osteoporosis and/or oligosaccharide. Assays can also be developed to screen
 CC compounds which are capable of altering the binding activity of the
 CC ligand affecting activation of the G protein-coupled receptor protein.
 XX
 SQ Sequence 22 AA;

Query Match 100.0%; Score 116; DB 18; Length 22;
 Best Local Similarity 100.0%; Pred. No. 1.4e-11;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYTGRIPIVGRF 20
 ||||||||||||||||
 DB 1 tpdinpawytgriprgrf 20

RESULT 9
 B10360

ID B10360 standard; peptide; 22 AA.
 AC B10360;
 DT 24-NOV-2000 (first entry)
 DE Rat oxytocin secretion promoting peptide SEQ ID NO: 23.
 XX
 KW Rat oxytocin secretion promoter; G protein-coupled receptor protein;
 KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;
 KW caesarean section; artificial fertilization; galactostasis; goat; pig;
 KW veterinary medicine; milk production.
 XX
 OS Rattus sp.
 XX
 PN W0200038704-A1.
 PD 06-JUL-2000.
 XX
 PF 22-DEC-1999; 99WO-JP07199.
 XX
 PR 25-DEC-1998; 98JP-0369585.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 PI Matsumoto H, Kitada C, Hinuma S;
 DR WPI: 2000-452298/39.
 XX
 PT Physiologically active polypeptide recognized as ligand by G
 PT protein-coupled receptor protein, for promoting secretion of oxytocin,
 PT as drugs for diseases relating to oxytocin secretion and in veterinary
 PT medicine
 XX
 PS Disclosure: Page 59; 72pp; Japanese.
 XX
 CC This invention describes a novel oxytocin secretion-regulating agent
 CC which contains a ligand peptide or its salt for the G protein-coupled
 CC receptor protein. It is useful in the form of drugs for ameliorating,
 CC preventing and treating diseases relating to oxytocin secretion e.g.,
 CC weak pains and atonic bleeding, before and after expulsion of placenta,
 CC uterine recovery failure, caesarean section, stoppage of artificial
 CC fertilization or galactostasis and is also applicable in veterinary
 CC medicine for promoting milk production in cow, goat and pig. This
 CC sequence represents a rat peptide which acts as an oxytocin secretion
 CC promoter.
 XX
 SQ Sequence 22 AA;

Query Match 100.0%; Score 116; DB 21; Length 22;
 Best Local Similarity 100.0%; Pred. No. 1.4e-11;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYTGRIPIVGRF 20
 ||||||||||||||||
 DB 1 tpdinpawytgriprgrf 20

RESULT 10
 ID W31384 standard; Peptide; 31 AA.
 AC W31384;

DT 06-APR-1998 (first entry)

DE Rat type G protein-coupled receptor ligand fragment 1.

KW G protein-coupled receptor; ligand binding; pharmaceutical;
 KW modulator; pituitary; central nervous system; pancreas; prophylactic;
 KW therapeutic agent.
 XX

```

OS   Rat sp.
PN   WO9724436-A2.
PD   10-JUL-1997.
PF   26-DEC-1996; 96WO-JP03821.
PR   18-SEP-1996; 96JP-0246573.
PR   28-DEC-1995; 95JP-0343371.
PR   15-MAR-1996; 96JP-0059419.
PR   12-AUG-1996; 96JP-0211805.
PA   (TAKE ) TAKEDA CHEM IND LTD.
P1   Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;
P1   Kawamata Y, Kitada C;
XX   WPI: 1997-363672/33.
DR   N-PSDB; V02421.
PT   Lysand peptide for G protein-coupled receptor - acts by modulating
PT   function in the central nervous system, pancreas and pituitary gland
XX   Claim 2; Page 179; 258pp; English.
XX
CC   This sequence represents a peptide fragment from a novel rat type
CC   1 ligand polypeptide corresponding to amino acid residues 22 to 52 of the
CC   sequence represented in W31383 and is used in an assay to monitor ligand
CC   binding to the G protein-coupled receptor protein. Pharmaceutical
CC   compositions containing this ligand may be used as a pituitary function
CC   modulator, a central nervous system modulator or a pancreatic function
CC   modulator. This ligand could have specific applications as a prophylactic
CC   or therapeutic agent for dementia, depression, hyperkinetic syndrome,
CC   disturbance of consciousness, anxiety syndrome, schizophrenia, trauma,
CC   growth hormone secretory disease, hyper- and hypoparathyroidism, diabetes,
CC   hypercholesterolemia, hyperglycidaemia, hyperprolactinaemia, osteoporosis,
CC   cancer, pancreatitis, renal disease, Turner's syndrome, neurosis, asthma,
CC   rheumatoid arthritis, spinal injury, transient brain ischaemia, epilepsy,
CC   amyotrophic lateral sclerosis, acute myocardial infarction, infertility,
CC   splanchnic lateral sclerolosis, bone fracture, trauma, atopic dermatitis,
CC   osteoporosis and/or oligosaccharia. Assays can also be developed to screen
CC   compounds which are capable of altering the binding activity of the
CC   ligand affecting activation of the G protein-coupled receptor protein.
XX
SQ   Sequence 31 AA;
XX
Query Match 100.0%; Score 116; DB 18; Length 31;
Best Local Similarity 100.0%; Pred. No. 2, 1e-11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
OY 1 PPDINPAWYTGIRPVGRF 20
DB 12 updinpawytgirtprgrf 31
XX
RESULT 11
ID W97233
AC W97233 standard; peptide; 31 AA.
XX
XX W97233:
XX
XX 06-MAY-1999 (first entry)
DE Rat type ligand polypeptide fragment.
XX
XX Rat type ligand; modulation; prolactin secretion;
XX G protein-coupled receptor; GPCR; hypovarianism; gonocyst cacosgenesis;
XX menopausal syndrome; euthyroid; hypometabolism; lactation;
XX pituitary adenomatosis; brain tumour; amenorrhoea; autoimmune disease;
XX prolactinoma; infertility; impotence; amenorrhoea; galactorrhoea;
XX acromegaly; Chiari-Frömmel syndrome; Argonz-del Castillo syndrome;

```

KV	Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;
KV	contraceptive; placental function; choriochorionoma; hydatid mole;
KV	irruption mole; abortion; unfertilized fetus; abnormal saccharometabolism;
KV	abnormal lipidmetabolism; oxytocin.
XX	
OS	Rattus sp.
XX	
PM	MO9858962-A1.
XX	
PD	30-DEC-1998.
XX	
PF	22-JUN-1998; 98WO-JP02765.
XX	
PR	23-JUN-1997; 97JP-0165437.
XX	
PA	(TAKE) TAKEDA CHEM IND LTD.
XX	
PI	Fuji R, Hinuma S, Kawamata Y, Matsumoto N;
XX	
DR	WPI, 1999-105614/09.
XX	
PT	Use of G protein-coupled receptor ligands - for modulating prolactin
PT	secretion or placental function, e.g. for treating menopause
PT	syndrome, tumours, autoimmune disease or abnormal pregnancy
XX	
PS	Claim 3; Page 153; 241pp; English.
XX	
CC	The present sequence represents a rat type ligand fragment. It
CC	is used in the course of the invention. The specification describes
CC	an agent for modulating prolactin secretion which comprises a
CC	ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)
CC	protein. The agents for promoting prolactin secretion can be used for
CC	treating or preventing hypovarianism, gonocyst carcinogenesis, menopause
CC	syndrome, eutryoid or hypometabolism. They can be used for promoting
CC	lactation in a domestic mammal and as an aphrodisiac. The agents for
CC	inhibiting prolactin secretion can be used for treating or preventing
CC	pituitary adenomatosis, brain tumour, amenorrhoea, autoimmune disease,
CC	prolactinoma, infertility, impotence, amenorrhoea, galactorrhea,
CC	acromegaly, Chlat-Frömmel syndrome, Argonz-del Castillo syndrome,
CC	Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dyszoospermia.
CC	The inhibitory agents can also be used as contraceptives. The agents for
CC	modulating placental function can be used for treating or preventing
CC	choriochorionoma, hydatid mole, irruption mole, abortion, unfertilized fetus,
CC	abnormal saccharometabolism, abnormal lipidmetabolism or oxytocin.
XX	
SO	Sequence 31 AA;
XX	
Query Match	100.0%; Score 116; DB 20; Length 31;
Best Local Similarity	100.0%; Pred. No. 2.1e-11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY	1 TPIDINPAWYTGIRPVGRF 20
DB	12 tpdinpawytgtrpvgrf 31
RESULT 12	
ID	W87614
AC	W87614 standard; Peptide; 31 AA.
XX	
XX	W87614;
XX	
XX	29-MAR-1999 (first entry)
XX	
DE	Rat 1992 ligand.
XX	
KV	1992 ligand; G protein coupled receptor; pituitary;
KV	prolactin releasing peptide; rat; dementia; breast cancer;
XX	therapy.
OS	Rattus sp.
XX	

EP867417-A2.
 30-DEC-1998.
 25-JUN-1998; 98EP-0111725.
 27-JUN-1997; 97JP-0172118.
 (TAKE) TAKEDA CHEM IND LTD.
 Moriya T, Nishimura O, Suenaga M, Tanaka Y;
 WPI, 1999-047884/05.
 Producing a 19p2 pituitary G protein receptor ligand - by cleavage of a fusion protein, useful for preventing and treating dementia, breast cancer, renal failure and autoimmune disease
 Claim 5; Page 34; 56pp; English.

This is the amino acid sequence of the rat pituitary G protein-coupled receptor ligand 19p2. A method suitable for commercial high-level production of 19p2 comprises expressing the ligand in host cells as a recombinant fusion protein e.g. with human basic fibroblast growth factor (see V83794-95) that has been modified to include an N-terminal cysteine residue. The ligand is released from the fusion by cyanylation followed by ammonolysis. 19p2 has prolactin secretion-stimulating and (at high doses) prolactin secretion-inhibiting properties. It can be used in the treatment and prevention of various diseases including: senile dementia, cerebrovascular dementia, and dementia associated with neurological disorders (e.g. Alzheimer's disease, Parkinson's disease, Pick's disease, Huntington's disease), infectious diseases (e.g. Creutzfeldt-Jakob's), endocrine or metabolic disease or intoxication by drugs, metal and organic compounds, tumorigenic diseases (e.g. brain tumour), traumatic diseases (e.g. chronic subarachnoid hemorrhage, and other types of dementia, depression, hyperactive child syndrome (microcephalopathy) and disturbance of consciousness. It is also useful for prevention and treatment of diseases associated with prolactin hypo and hypersecretion respectively, including: hyperprolactinaemia, pituitary adenoma, breast cancer, infertility, impotence and autoimmune disease (hypersecretion disorders), and seminal vesicle hypoplasia, osteoporosis, menopausal syndrome and renal failure (hyposecretion disorders). The 19p2 polypeptide/amide is also useful as a test reagent for study of the prolactin secretory function or as a lactagogue in mammalian farm animals.

Sequence 31 AA;

Query Match 100.0%; Score 116; DB 20; Length 31;
 Best Local Similarity 100.0%; Pred. No. 2.1e-11;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 TPDINPAMYTGRIPIRGF 20
 12 tpdinpamyltgrlprvgrf 31

RESULT 13
 W95173
 W95173 standard; peptide; 31 AA.

10-MAR-1999 (first entry)
 Murine pituitary-derived ligand mature polypeptide sequence.
 Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor; GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas;

tissue; screen; therapeutic; binding; senile dementia; ligand; murine; Alzheimer's disease; Parkinson's disease; Huntington's disease; drug; Creutzfeldt-Jakob disease; poisoning; schizophrenia; growth hormone; secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor; gene therapy; transgenic animal.

Mus sp.

W09849295-A1.

05-NOV-1998.

27-APR-1998; 98WO-JP01923.

28-APR-1997; 97JP-0109974.

(TAKE) TAKEDA CHEM IND LTD.

Fukusumi S, Hinuma S;

WPI, 1999-009423/01.

New polypeptide ligand for orphan G protein coupled receptors - used for treating disorders of central nervous system, pituitary and pancreas, and for drug screening

Disclosure: Page 134; 206pp; English.

This represents the matured murine pituitary-derived ligand polypeptide sequence. The polypeptide is a ligand for the G-protein coupled orphan receptor designated GPR10 (human) or UHR-1 (rat). Cells transformed with a vector containing the ligand polypeptide encoding DNA are used to produce a recombinant ligand polypeptide. The ligand polypeptide, and its fragments, modulate function of the pituitary, central nervous system, pancreas and other tissues and can be used to screen for agents that modulate binding of the polypeptide to the receptor; to quantify the amount of receptor in a sample and to raise antibodies. They may also be used therapeutically, e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's diseases; Creutzfeldt-Jakob disease; poisoning by heavy metals or drugs; diabetes; schizophrenia; disorders of growth hormone secretion; cancer; rheumatoid arthritis, epilepsy and many others, also to improve post-operative nutritional status and as vasopressor. Transgenic animals carrying the ligand polypeptide encoding DNA or its mutein are used to study the function of the polypeptide-expressing genes, as models of disease, for drug screening and as source of cell lines. The ligand polypeptide DNA is used as a source of probes and primers; to identify related sequences; in receptor-binding assays; for production of Ab and antisera; in drug development; for gene therapy and to develop transgenic animals.

Sequence 31 AA;

Query Match 100.0%; Score 116; DB 20; Length 31;
 Best Local Similarity 100.0%; Pred. No. 2.1e-11;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 TPDINPAMYTGRIPIRGF 20
 12 tpdinpamyltgrlprvgrf 31

RESULT 14
 W95174
 W95174 standard; protein; 31 AA.

10-MAR-1999 (first entry)
 Murine pituitary-derived ligand polypeptide antigenic epitope.
 Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;

KM GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas;
 KM tissue; screen; therapeutic; binding; senile dementia; ligand; murine;
 KM Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;
 KM Creutzfeldt-Jakob disease; poisoning; schizophrenia; growth hormone;
 KM secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;
 KM gene therapy; transgenic animal; epitope.

XX Mus sp.

XX WO9849295-A1.

XX 05-NOV-1998.

XX 27-APR-1998; 98WO-JP01923.

XX 28-APR-1997; 97JP-0109974.

XX (TAKE) TAKEDA CHEM IND LTD.

XX Fukusumi S, Hinuma S;

XX WPI; 1999-009423/01.

XX New polypeptide ligand for orphan G protein coupled receptors - used
 PT for treating disorders of central nervous system, pituitary and
 PT pancreas, and for drug screening

XX Disclosure; Page 26; 206pp; English.

XX The invention relates to a murine pituitary-derived ligand polypeptide
 CC which is a ligand for the G-protein coupled orphan receptor designated
 CC GPR10 (human) or UHR-1 (rat). Cells transformed with a vector containing
 CC the ligand polypeptide encoding DNA are used to produce a recombinant
 CC ligand polypeptide. The ligand polypeptide, and its fragments, modulate
 CC function of the pituitary, central nervous system, pancreas and other
 CC tissues and can be used to screen for agents that modulate binding of the
 CC polypeptide to the receptor; to quantify the amount of receptor in a
 CC sample and to raise antibodies. They may also be used therapeutically,
 CC e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's
 CC diseases; Creutzfeldt-Jakob disease; poisoning by heavy metals or drugs;
 CC diabetes; schizophrenia; disorders of growth hormone secretion; cancer;
 CC rheumatoid arthritis; epilepsy and many others, also to improve post-
 CC operative nutritional status and as vasopressor. Transgenic animals
 CC carrying the ligand polypeptide encoding DNA or its mutin are used to
 CC study the function of the polypeptide-expressing genes, as models of
 CC disease, for drug screening and as source of probes and primers; to identify
 CC polypeptide DNA is used as a source of probes and primers; to identify
 CC related sequences; in receptor-binding assays; for production of Ab and
 CC antisera; in drug development; for gene therapy and to develop transgenic
 CC animals. Sequences W95174 to W95178 represent antigenic epitopes which
 CC can be used for the preparation of anti-ligand polypeptide antibody.

XX Sequence 31 AA:

Query Match 100.0%; Score 116; DB 20; Length 31;

Best Local Similarity 100.0%; Pred. No. 2.1e-11;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TPDIINPAWYTGRIIPVGRF 20

DB 12 tpdinpwytgrtprvgrf 31

RESUME 15

ID B10355 standard; peptide; 31 AA.

AC B10355;

DT 24-NOV-2000 (first entry)

DE Rat oxytocin secretion promoting peptide SEQ ID NO: 18.

XX Rat; oxytocin secretion promoter; G protein-coupled receptor protein;
 KM treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;
 KM caesarean section; artificial fertilization; galactostasis; goat; pig;
 KM veterinary medicine; milk production.

XX Rattus sp.

XX WO200038704-A1.

XX 06-JUL-2000.

XX 22-DEC-1999; 99WO-JP07199.

XX 25-DEC-1998; 98JP-0369585.

XX (TAKE) TAKEDA CHEM IND LTD.

XX Matsumoto H, Kikada C, Hinuma S;

XX WPI; 2000-452298/39.

XX Physiologically active polypeptide recognized as ligand by G
 PT protein-coupled receptor protein, for promoting secretion of oxytocin,
 PT as drugs for diseases relating to oxytocin secretion and in veterinary
 PT medicine

XX Claim 3; Page 57; 72pp; Japanese.

XX This invention describes a novel oxytocin secretion-regulating agent
 CC which contains a ligand peptide or its salt for the G protein-coupled
 CC receptor protein. It is useful in the form of drugs for ameliorating,
 CC preventing and treating diseases relating to oxytocin secretion e.g.
 CC weak pains and atonic bleeding, before and after expulsion of placenta,
 CC uterine recovery failure, caesarean section, stoppage of artificial
 CC fertilization or galactostasis and is also applicable in veterinary
 CC medicine for promoting milk production in cow, goat and pig. This
 CC sequence represents a rat peptide which acts as an oxytocin secretion
 CC promoter.

XX Sequence 31 AA:

Query Match 100.0%; Score 116; DB 21; Length 31;

Best Local Similarity 100.0%; Pred. No. 2.1e-11;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TPDIINPAWYTGRIIPVGRF 20

DB 12 tpdinpwytgrtprvgrf 31

Search completed: April 17, 2001, 15:38:42
 Job time: 349 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: April 17, 2001, 15:39:48 ; Search time 61.54 Seconds
(without alignments)
6.243 Million cell updates/sec

Title: US-09-446-543a-50

Sequence: I TPDINPAWYNGRCIRPVGRF 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Issued Patents_Aa:*

- 1: /cgn2.6/ptodata/2/1aa/5a.COMB.pep:*
- 2: /cgn2.6/ptodata/2/1aa/5b.COMB.pep:*
- 3: /cgn2.6/ptodata/2/1aa/6a.COMB.pep:*
- 4: /cgn2.6/ptodata/2/1aa/6b.COMB.pep:*
- 5: /cgn2.6/ptodata/2/1aa/PCITUS.COMB.pep:*
- 6: /cgn2.6/ptodata/2/1aa/Dackillest.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	116	100.0	20	US-09-105-678A-40	Sequence 40, Appl
2	116	100.0	21	US-09-105-678A-41	Sequence 41, Appl
3	116	100.0	22	US-09-105-678A-42	Sequence 42, Appl
4	116	100.0	31	US-09-105-678A-8	Sequence 8, Appl
5	116	100.0	31	US-09-105-678A-37	Sequence 37, Appl
6	116	100.0	31	US-09-172-353-4	Sequence 4, Appl
7	116	100.0	33	US-09-105-678A-38	Sequence 38, Appl
8	116	100.0	33	US-09-105-678A-39	Sequence 39, Appl
9	111	95.7	20	US-09-105-678A-34	Sequence 34, Appl
10	111	95.7	21	US-09-105-678A-35	Sequence 35, Appl
11	111	95.7	22	US-09-105-678A-36	Sequence 36, Appl
12	111	95.7	31	US-09-105-678A-7	Sequence 7, Appl
13	111	95.7	31	US-09-105-678A-31	Sequence 31, Appl
14	111	95.7	32	US-09-105-678A-32	Sequence 32, Appl
15	111	95.7	33	US-09-105-678A-33	Sequence 33, Appl
16	105	90.5	19	US-09-105-678A-30	Sequence 30, Appl
17	105	90.5	20	US-09-105-678A-46	Sequence 46, Appl
18	105	90.5	21	US-09-105-678A-47	Sequence 47, Appl
19	105	90.5	22	US-09-105-678A-48	Sequence 48, Appl
20	105	90.5	31	US-09-105-678A-9	Sequence 9, Appl
21	105	90.5	31	US-09-105-678A-43	Sequence 43, Appl
22	105	90.5	32	US-09-105-678A-44	Sequence 44, Appl
23	104	89.7	33	US-09-105-678A-45	Sequence 45, Appl
24	104	89.7	21	US-09-105-678A-28	Sequence 28, Appl
25	100	86.2	29	US-09-105-678A-29	Sequence 29, Appl
26	50	43.1	349	US-08-118-270-71	Sequence 71, Appl
27	50	43.1	349	PCT-US93-08528-71	Sequence 71, Appl

28	49	42.2	774	3	US-08-902-632-2	Sequence 2, Appl
29	49	42.2	774	3	US-09-073-354-1	Sequence 1, Appl
30	49	42.2	774	3	US-08-656-005A-1	Sequence 1, Appl
31	49	42.2	774	4	US-09-073-259-1	Sequence 1, Appl
32	49	42.2	774	4	US-09-363-095-1	Sequence 1, Appl
33	46	39.7	330	2	US-08-815-176-1	Sequence 1, Appl
34	46	39.7	555	2	US-08-982-232-14	Sequence 14, Appl
35	46	39.7	555	2	US-08-982-232-14	Sequence 14, Appl
36	45	38.8	778	1	US-08-906-925-4	Sequence 4, Appl
37	45	38.8	779	1	US-08-375-134-12	Sequence 12, Appl
38	45	38.8	779	5	PCT-US95-15263-12	Sequence 12, Appl
39	45	38.8	940	4	US-09-078-347A-1	Sequence 1, Appl
40	44	37.9	1019	1	US-08-271-364A-7	Sequence 7, Appl
41	44	37.9	1019	2	US-08-222-715B-26	Sequence 26, Appl
42	44	37.1	274	4	US-09-141-821-3	Sequence 3, Appl
43	41	35.3	70	4	US-08-513-974B-34	Sequence 34, Appl
44	41	35.3	70	4	US-08-513-974B-317	Sequence 317, App
45	41	35.3	141	4	US-08-513-974B-320	Sequence 320, App

ALIGNMENTS

RESULT 1
US-09-105-678A-40
Sequence 40, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172116/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-40

Query Match 100.0%; Score 116; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.8e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TPDINPAWYTGIRPVGRF 20
|||||
Db 1 TPDINPAWYTGIRPVGRF 20

RESULT 2

US-09-105-678A-41
; Sequence 41, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,678A
; FILING DATE: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-105-678A-41

Query Match 100.0%; Score 116; DB 3; Length 21;

Best Local Similarity 100.0%; Pred. No. 5, Je-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TPDINPAWYTGIRPVGRF 20
|||||
Db 1 TPDINPAWYTGIRPVGRF 20

RESULT 3

US-09-105-678A-42
; Sequence 42, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-42

Query Match 100.0%; Score 116; DB 3; Length 22;
Best Local Similarity 100.0%; Pred. No. 5, Je-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TPDINPAWYTGIRPVGRF 20
|||||
Db 1 TPDINPAWYTGIRPVGRF 20

RESULT 4

US-09-105-678A-8
; Sequence 8, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,678A
; FILING DATE: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.

REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-8

Query Match 100.0%; Score 116; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 7.8e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAMWTGGRIPVGRF 20
|||||
Db 12 TPDINPAMWTGGRIPVGRF 31

RESULT 5
US-09-105-678A-37
Sequence 37, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-37

Query Match 100.0%; Score 116; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 7.8e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAMWTGGRIPVGRF 20
|||||
Db 12 TPDINPAMWTGGRIPVGRF 31

RESULT 6
US-09-172-353-4
Sequence 4, Application US/09172353
Patent No. 6197530
GENERAL INFORMATION:
APPLICANT: Stricker-Kongra, Alain
APPLICANT: Gu, Wei
TITLE OF INVENTION: GPRI0 AS A TARGET FOR IDENTIFYING WEIGHT MODULATING COMPOUNDS
FILE REFERENCE: 07334/102001
CURRENT APPLICATION NUMBER: US/09/172,353
CURRENT FILING DATE: 1998-10-14
NUMBER OF SEQ ID NOS: 7
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
LENGTH: 31
TYPE: PRT
ORGANISM: Mus musculus
US-09-172-353-4

Query Match 100.0%; Score 116; DB 4; Length 31;
Best Local Similarity 100.0%; Pred. No. 7.8e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAMWTGGRIPVGRF 20
|||||
Db 12 TPDINPAMWTGGRIPVGRF 31

RESULT 7
US-09-105-678A-38
Sequence 38, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:

LENGTH: 32 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-38

Query Match 100.0%; Score 116; DB 3; Length 32;
Best Local Similarity 100.0%; Pred. No. 8,1e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TPDINPAMWTGIRGIRPVGRF 20
|||||
DB 12 TPDINPAMWTGIRGIRPVGRF 31

RESULT 8
US-09-105-678A-39
Sequence 39, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-39

Query Match 100.0%; Score 116; DB 3; Length 33;
Best Local Similarity 100.0%; Pred. No. 8,4e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TPDINPAMWTGIRGIRPVGRF 20
|||||
DB 12 TPDINPAMWTGIRGIRPVGRF 31

RESULT 9

US-09-105-678A-34
Sequence 34, Application US/09105678A
Patent No. 6103882

GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-34

Query Match 95.7%; Score 111; DB 3; Length 20;
Best Local Similarity 95.0%; Pred. No. 2,9e-11;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TPDINPAMWTGIRGIRPVGRF 20
|||||
DB 1 TPDINPAMWTGIRGIRPVGRF 20

RESULT 10
US-09-105-678A-35
Sequence 35, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-35

Query Match 95.7%; Score 111; DB 3; Length 21;
Best Local Similarity 95.0%; Pred. No. 3.1e-11;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TPDINPAMVYTGRCIRPVGRF 20
|||||
Db 1 TPDINPAMVYTGRCIRPVGRF 20

RESULT 11
US-09-105-678A-36
Sequence 36, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 36:

SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-36

Query Match 95.7%; Score 111; DB 3; Length 22;
Best Local Similarity 95.0%; Pred. No. 3.2e-11;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TPDINPAMVYTGRCIRPVGRF 20
|||||
Db 1 TPDINPAMVYTGRCIRPVGRF 20

RESULT 12
US-09-105-678A-7
Sequence 7, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-7

Query Match 95.7%; Score 111; DB 3; Length 31;
Best Local Similarity 95.0%; Pred. No. 4.7e-11;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TPDINPAMVYTGRCIRPVGRF 20
|||||
Db 12 TPDINPAMVYTGRCIRPVGRF 31

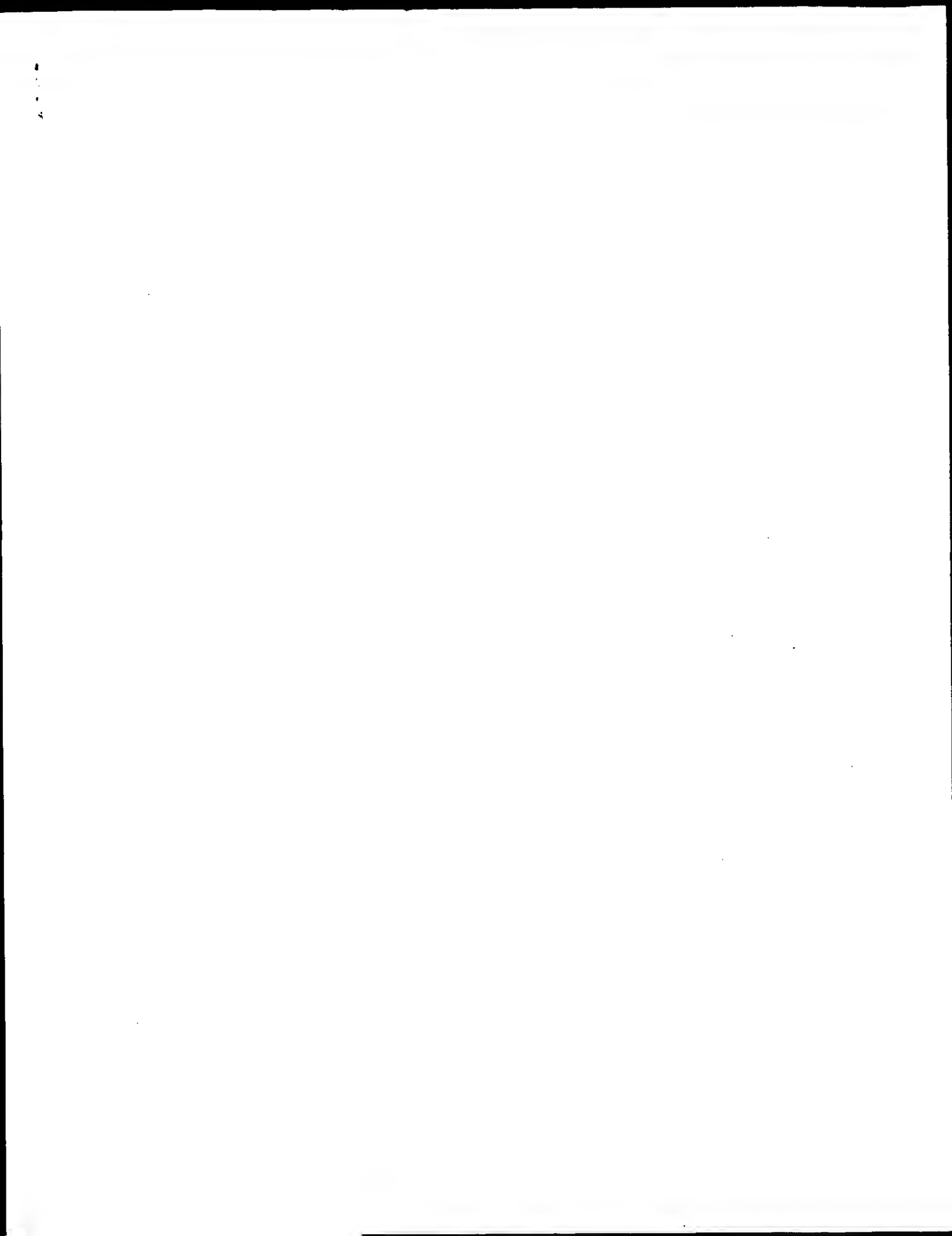
RESULT 13
US-09-105-678A-31
; Sequence 31, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,678A
; FILING DATE: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conliln, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; STRANDEDNESS: linear
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-105-678A-31
Query Match 95.7%; Score 111; DB 3; Length 31;
Best Local Similarity 95.0%; Pred. No. 4,7e-11;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 TPDINPAWYTGIRPVGRF 20
DB 12 TPDINPAWYTGIRPVGRF 31
RESULT 14
US-09-105-678A-32
; Sequence 32, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,678A
; FILING DATE: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conliln, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-105-678A-32
Query Match 95.7%; Score 111; DB 3; Length 32;
Best Local Similarity 95.0%; Pred. No. 4,9e-11;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 TPDINPAWYTGIRPVGRF 20
DB 12 TPDINPAWYTGIRPVGRF 31
RESULT 15
US-09-105-678A-33
; Sequence 33, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,678A
; FILING DATE: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conliln, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440

; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-105-678A-33

Query Match 95.7%; Score 111; DB 3; Length 33;
Best Local Similarity 95.0%; Pred. No. 5, 1e-11;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 TPDINPAMWTGGRGIRPVGRF 20
||||||| |||||
DB 12 TPDINPAMWTGGRGIRPVGRF 31

Search completed: April 17, 2001, 15:39:48
Job time: 317 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 17, 2001, 15:48:48 ; Search time 39.1 Seconds

Title: US-09-446-543A-50

Perfect score:

Sequence: 1 TPDINPAMYTGGRIPVGRF 20

Scoring table:

Gapop 10.0 , Gapext 0.5

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08

Maximum	Match	100%
100%	100%	100%

Listing first 45 summaries

Database : SWISSProt_39:★

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	116	100.0	93	1	PRRR_RAT	P81278 rattus norv
2	111	95.7	98	1	PRRR_BOVIN	P81264 bos taurus
3	105	90.5	87	1	PRRR_HUMAN	P81277 homo sapien
4	51	44.0	962	1	UVRA_METTH	O26543 methanobact
5	47	40.5	453	1	TBB2_GBOCN	P32925 geotrichum
6	47	40.5	719	1	ARP_YEAST	P32925 geotrichum
7	47	40.5	972	1	UVRA_MYCTU	P32170 saccharomyce
8	47	40.5	973	1	UVRA_RHIME	P94972 mycobacteri
9	46.5	40.1	652	1	TEPT_CLOPE	P56899 rhlzobium m
10	46	39.7	320	1	GRP2_MOUSE	O46306 clostridium
11	46	39.7	330	1	GRP2_HUMAN	O89190 m_grib2-reia
12	46	39.7	943	1	UVRA_HAEIN	O75921 n_grib2-reia
13	45	38.8	482	1	R167_YEAST	P44410 haemophilus
14	45	38.8	676	1	EXL1_HUMAN	P39743 saccharomyce
15	45	38.8	773	1	DPOL_THEGO	O92835 homo saplen
16	45	38.8	798	1	UNR_RAT	P56889 thermococcu
17	45	38.8	940	1	UVRA_BOOLI	P18395 rattus norv
18	45	38.8	941	1	UVRA_SALTY	P07871 escherichia
19	45	38.8	953	1	UVRA_RICPR	P74734 salmonella
20	45	38.8	1523	1	DPOL_THIEFM	O92933 rickettsia
21	44	37.9	179	1	RK6_GUTTH	P74918 thermococcu
22	44	37.9	184	1	RL6_THEMA	O46908 guillardia
23	44	37.9	771	1	DPOL_PVRAB	O92944 thermotoga
24	44	37.9	771	1	DPOL_PVRSE	P77916 pyrococcus
25	44	37.9	952	1	UVRA_THETH	P77932 pyrococcus
26	44	37.9	970	1	UVRA_SYNY3	O56442 thermus aqu
27	44	37.9	1014	1	UVRA_SYRSCO	P73412 synchocyst
28	44	37.9	1312	1	DPOL_PYRSD	O92507 streptomyce
29	43.5	37.5	1680	1	FUR2_DROME	O51334 pyrococcus
30	43	37.1	241	1	YC3D_YEAST	P30432 dirosophila
31	43	37.1	272	1	LCB3_YEAST	P53381 saccharomyce
32	43	37.1	476	1	YAAV_ECOLI	O41160 roblinia pse
33	43	37.1	569	1	UVRA_VTST	P30143 escherichia
						O08518 vltresocill

34	43	37.1	950	1	UVPA_METIO	Q50968	neisseria 9
35	43	360	1	1	UVPA_TREPA	083527	treponema p
36	43	37.1	1282	1	TP2M_DIPIA	P03550	diclofenac p
37	42.5	36.6	486	1	DMPG_PSESP	P190059	pseudomonas
38	42	36.2	247	1	MERS_ARATH	P24806	arabidopsist
39	42	36.2	272	1	TR42_DROVI	O02008	drosophila
40	42	36.2	424	1	SAMB_SALTY	P23882	salmonella
41	42	36.2	926	1	UVPA_AQUAE	066911	aquifex aco
42	42	36.2	1235	1	DPOL_PYRRO	059610	pyrococcus
43	41	36.2	41	1	UVPA_PPMIS	Q05329	mycobacteri
44	41	35.3	101	1	VG30_PPMIS	P57416	buchnera ap
45	41	35.3	264	1	Y355_BUCMA	058172	methanococc
46	41	35.3	342	1	Y762_METCA		

ALIGNMENTS

	RESULT	1
PRRP_RAT	PRRP_RAT	STANDARD; PRT; 83 AA.
ID	PRRP_RAT	
AC	P81278;	
DT	30-MAY-2000 (Rel. 39, Created)	
DT	30-MAY-2000 (Rel. 39, Last sequence update)	
DT	30-MAY-2000 (Rel. 39, Last annotation update)	
DE	PROLACTIN-RELEASING PEPTIDE PRECURSOR (PRP) (PROLACTIN-RELEASING HORMONE) [CONTAINS: PROLACTIN-RELEASING PEPTIDE PRP31; PROLACTIN-RELEASING PEPTIDE PRRP20].	
GN	PRP.	
OS	Rattus norvegicus (Rat).	
OC	Eumetazoa; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
CC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	
OX	NCBI_TaxID=10116;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
PC	TISSUE=Brain:	
RX	MEDLINE=98268781; PubMed=9607765;	
RA	Hinuma S., Nabata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,	
SA	Kilada S.C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,	
RA	Kurokawa T., Nishimura O., Onda H., Fujiwara M.;	
RT	"A prolactin-releasing peptide in the brain.";	
RL	Nature 393:272-276(1998).	
CC	-I- FUNCTION: STIMULATES PROLACTIN (PR) RELEASE AND REGULATES THE EXPRESSION OF PROLACTIN THROUGH ITS RECEPTOR HGRC3. MAY STIMULATE LACTOTROPHS DIRECTLY TO SECRETE PR.	
CC	-I- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch .	
CC	--	
DR	EMBL; AB015418; BAA29026.1; "	
FT	Hormone; Amidation; signal; Cleavage area pair of basic residues.	
FT	SIGNAL	1 21 BY SIMILARITY
FT	PEPTIDE	22 52 PROLACTIN-RELEASING PEPTIDE PRP31.
FT	PEPTIDE	33 52 PROLACTIN-RELEASING PEPTIDE PRRP20.
FT	MOD_RES	52 52 AMIDATION (G-53 PROVIDE AMIDE GROUP).
SQ	SEQUENCE	83 AAP: 9215 MW; DOCT5264EEBEAF29 CRC64;

Query Match	100.0%;	Score 116;	DB 1;	Length 83;
Best Local Similarity	100.0%;	Pred. No. 5.8e-11;		
Matches	20;	Conservative	0;	Mismatches 0;
				Indels 0;
				Gaps 0;

```

QY      1 TPDINPAWYTGGRIPVGRF 20
        |||||
Db      33 TPDINPAWYTGGRIPVGRF 52

```

RESULT 2

```

PRRP_BOVIN          STANDARD;          PRT;          98 AA.
ID   PRRP_BOVIN          P81264;
DI   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DE   30-MAY-2000 (Rel. 39, Last annotation update)
DE   PROLACTIN-RELEASING PEPTIDE PRECURSOR (PRRP) (PROLACTIN-RELEASING
DE   HORMONE) [CONTAINS: PROLACTIN-RELEASING PEPTIDE PRRP31; PROLACTIN-
DE   RELEASING PEPTIDE PRRP20].
GN   PRR.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 23-52.
RC   TISSUE=Brain;
RX   MEDLINE=98268781; PubMed=9607765;
RA   Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,
RA   Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,
RA   Kurokawa T., Nishimura O., Onda H., Fujino M.;
RT   "A prolactin-releasing peptide in the brain.";
RL   Nature 393:272-276(1998).
CC   -1- FUNCTION: STIMULATES PROLACTIN (PRL) RELEASE AND REGULATES THE
CC   EXPRESSION OF PROLACTIN THROUGH ITS RECEPTOR HGR3. MAY STIMULATE
CC   LACTOTROPHS DIRECTLY TO SECRETE PRL.
CC   -1- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL; AB015417; BAA29025.1; -
KM   Hormone; Amidation; Signal; Cleavage on pair of basic residues.
FT   SIGNAL          1      22
FT   PEPTIDE         23      53
FT   PEPTIDE         33      53
FT   MOD_RES         53      53
SQ   SEQUENCE 98 AA; 10544 MW; 08AC35A13B0FA908 CRC64;
Query Match          95.78; Score 111; DB 1; Length 98;
Best Local Similarity 95.08; Pred. No. 3,8e-10;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY   1 TPDIINPAMWTGRLRPVGRF 20
    |||||  |||||
DB   34 TPDIINPAMWTGRLRPVGRF 53

```

```

RX   MEDLINE=98268781; PubMed=9607765;
RA   Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,
RA   Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,
RA   Kurokawa T., Nishimura O., Onda H., Fujino M.;
RT   "A prolactin-releasing peptide in the brain.";
RL   Nature 393:272-276(1998).
CC   -1- FUNCTION: STIMULATES PROLACTIN (PRL) RELEASE AND REGULATES THE
CC   EXPRESSION OF PROLACTIN THROUGH ITS RECEPTOR HGR3. MAY STIMULATE
CC   LACTOTROPHS DIRECTLY TO SECRETE PRL.
CC   -1- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL; AB015419; BAA29027.1; -
KM   MIM: 602663; -
FT   SIGNAL          1      22
FT   PEPTIDE         23      53
FT   PEPTIDE         34      53
FT   MOD_RES         53      53
SQ   SEQUENCE 87 AA; 9639 MW; 229A2F3F50CF981B CRC64;
Query Match          90.5%; Score 105; DB 1; Length 87;
Best Local Similarity 90.0%; Pred. No. 2,8e-09;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY   1 TPDIINPAMWTGRLRPVGRF 20
    |||||  |||||
DB   34 TPDIINPAMWTGRLRPVGRF 53

```

```

RESULT 3
PRRP_HUMAN          STANDARD;          PRT;          87 AA.
ID   PRRP_HUMAN          P8127;
DI   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DE   30-MAY-2000 (Rel. 39, Last annotation update)
DE   PROLACTIN-RELEASING PEPTIDE PRECURSOR (PRRP) (PROLACTIN-RELEASING
DE   HORMONE) [CONTAINS: PROLACTIN-RELEASING PEPTIDE PRRP31; PROLACTIN-
DE   RELEASING PEPTIDE PRRP20].
GN   PRR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;

```

```

RESULT 4
ID   UVRA_METTH          STANDARD;          PRT;          962 AA.
DI   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DE   30-MAY-2000 (Rel. 39, Last annotation update)
DE   EXCINUCLEASE ABC SUBUNIT A.
GN   UVRA OR MTH43.
OS   Methanobacterium thermoautotrophicum.
OC   Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
OC   Methanothermobacter.
OX   NCBI_TaxID=145262;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DELTA H;
RX   MEDLINE=98037514; PubMed=9371463;
RA   Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA   Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert R.,
RA   Harrison D., Hoang L., Keagle P., Lum W., Pothier B., Qiu D.,
RA   Spadofora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA   Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA   McDougall S., Shiner G., Goyal A., Pietrowski S., Church G.M.,
RA   Daniels C.J., Mao J.-I., Rice P., Nolling J., Reeve J.N.;
RT   "Complete genome sequence of Methanobacterium thermoautotrophicum
RT   delta: functional analysis and comparative genomics.";
RL   J. Bacteriol. 179:7135-7155(1997).
CC   -1- FUNCTION: THE ABC EXCISION NUCLEASE IS A DNA REPAIR ENZYME THAT
CC   CATALYZES THE EXCISION REACTION OF UV-DAMAGED NUCLEOTIDE SEGMENTS
CC   PRODUCING OLIGOMERS HAVING THE MODIFIED BASE(S). UVRA IS AN ATPASE
CC   AND A DNA-BINDING PROTEIN THAT PREFERENTIALLY BINDS SINGLE-
CC   STRANDED OR UV-IRRADIATED DOUBLE-STRANDED DNA (BY SIMILARITY).
CC   -1- SUBUNIT: CONSISTS OF THREE SUBUNITS; UVRA, UVRB AND UVRC.
CC   -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC   -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY

```


Best Local Similarity 58.3%; Pred. No. 12;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 7 AMYTGIRPVG 18
DB 244 SMFTGYGVPRVG 255

RESULT 7
ID VURA_MYCTU STANDARD; PRT; 972 AA.
AC P94972;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE EXCINUCLEASE ABC SUBUNIT A.
GN VURA OR RV1638 OR MYC05H11.02.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Norris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Horsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R., Sulston J.E.,
RA Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
CC -1- FUNCTION: THE ABC EXCISION NUCLEASE IS A DNA REPAIR ENZYME THAT
CC CATALYZES THE EXCISION REACTION OF UV-DAMAGED NUCLEOTIDE SEGMENTS
CC PRODUCING OLIGOMERS HAVING THE MODIFIED BASE(S). VURA IS AN ATPASE
CC AND A DNA-BINDING PROTEIN THAT PREFERENTIALLY BINDS SINGLE-
CC STRANDED OR UV-IRRADIATED DOUBLE-STRANDED DNA (BY SIMILARITY).
CC -1- SUBUNIT: CONSISTS OF THREE SUBUNITS: VURA, UVRB AND UVRC.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC (ABC TRANSPORTERS). CONTAINS TWO ABC DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC
CC EMBL: Z85982; CAB06633.1; -
DR Tuberculolact; RV1638; -
DR InterPro: IPR001617; -
DR Pfam: PF00005; ABC_tran.2.
DR PROSITE: PS00211; ABC_TRANSPORTER.2.
KW SOS response; Excision nuclease; DNA repair; ATP-binding; Repeat;
KW DNA-binding; Zinc-finger.
FT NP_BIND 32 39 ATP (POTENTIAL).
FT NP_BIND 654 661 ATP (POTENTIAL).
FT ZN_FING 257 285 C4-TYPE (ATYPICAL).
FT ZN_FING 753 779 C4-TYPE.
SQ SEQUENCE 972 AA; 106131 MW; 8937A764E592D981 CRC64;

DB 709 TPRSNAFTYGVFKIRTLFAATTEAKRGYQP-GRF 744

RESULT 8
ID VURA_RHIME STANDARD; PRT; 973 AA.
AC P56899;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE EXCINUCLEASE ABC SUBUNIT A.
GN VURA.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-1021;
RX Galliard F., Capela D., Hubler-Barloy F., Gatius M., Batut J.,
RA Boistard P., Gonzy J., Kahn D., Thebaud P., Goffeau A.,
RA Purnelle B., Pohl T., Botne G., Schneider S., Portetelle D.,
RA Vandenbol M., Puehler A., Becker A., Weidner S.;
RL Submitted (MAY-2000) to the SWISS-PROT data bank.
RN (2)
RP SEQUENCE OF 1-140 FROM N.A.
RC STRAIN-2021;
RX MEDLINE=99430868; PubMed=10503543;
RA Tapias A., Barde J.;
RT "Regulation of divergent transcription from the vura-sab promoters in
RT Sinorhizobium meliloti."
RL Mol. Gen. Genet. 262:121-130(1999).
CC -1- FUNCTION: THE ABC EXCISION NUCLEASE IS A DNA REPAIR ENZYME THAT
CC CATALYZES THE EXCISION REACTION OF UV-DAMAGED NUCLEOTIDE SEGMENTS
CC PRODUCING OLIGOMERS HAVING THE MODIFIED BASE(S). VURA IS AN ATPASE
CC AND A DNA-BINDING PROTEIN THAT PREFERENTIALLY BINDS SINGLE-
CC STRANDED OR UV-IRRADIATED DOUBLE-STRANDED DNA (BY SIMILARITY).
CC -1- SUBUNIT: CONSISTS OF THREE SUBUNITS: VURA, UVRB AND UVRC.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC (ABC TRANSPORTERS). CONTAINS TWO ABC DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC
CC EMBL: AF125162; AAF03210.1; -
DR InterPro: IPR001617; -
DR PROSITE: PS00211; ABC_TRANSPORTER. PARTIAL.
KW SOS response; Excision nuclease; DNA repair; ATP-binding; Repeat;
KW DNA-binding; Zinc-finger.
FT NP_BIND 34 41 ATP (POTENTIAL).
FT NP_BIND 662 669 ATP (POTENTIAL).
FT ZN_FING 761 787 C4-TYPE.
FT ZN_FING 19 19 G -> A (IN REF. 2).
FT CONFLICT 19 19 F -> S (IN REF. 2).
SQ SEQUENCE 973 AA; 107191 MW; 3E1A8B14527A47FE CRC64;

Query Match 40.5%; Score 47; DB 1; Length 973;
Best Local Similarity 37.8%; Pred. No. 16;
Matches 14; Conservative 1; Mismatches 4; Indels 18; Gaps 2;

QY 1 TPDINPAMYTG-----RGIRPVGR 20
DB 717 TPRSNAFTYGVFKIRTLFAATTEAKRGYQP-GRF 752

QY 1 TPDINPAMYTG-----RGIRPVGR 20

RESULT 9
 ID TEMP CLOPE STANDARD: PRT: 652 AA.
 AC 046306;
 DT 15-DEC-1998 (rel. 37, Created)
 DT 15-DEC-1998 (rel. 37, Last sequence update)
 DT 15-DEC-1998 (rel. 37, Last annotation update)
 DE TETRACYCLINE RESISTANCE PROTEIN TETP (TETB(P)).
 GN TETP OR TETB(P).
 OS Clostridium perfringens.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OC Clostridium.
 OC NCBI_TaxID=1502;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CM92;
 RX MEDLINE=94224158; PubMed=8170402;
 RA Sloan J., McMurry L.M., Lytras D., Levy S.B., Rood J.I.;
 RT "The clostridium perfringens tet P determinant comprises two
 RT overlapping genes: tetA(P), which mediates active tetracycline
 RT efflux, and tetB(P), which is related to the ribosomal protection
 RT family of tetracycline-resistance determinants.";
 RL Mol. Microbiol. 11:403-415(1994).
 CC -1- FUNCTION: ABOLISH THE INHIBITORY EFFECT OF TETRACYCLIN ON PROTEIN
 CC SYNTHESIS BY A NON-COVALENT MODIFICATION OF THE RIBOSOMES.
 CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
 CC TETP/TETO SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L20800; AAA20117.1; -;
 CC DR InterPro: IPR000640; -;
 CC DR InterPro: IPR000795; -;
 CC DR InterPro: IPR002127; -;
 CC DR Pfam: PF00679; ERF_C; 1.
 CC DR Pfam: PF00009; GTP_EFTU; 1.
 CC DR PRINTS: PR00315; ELONGATNPFCT.
 CC DR PRINTS: PR01037; TCRTETOM.
 CC DR PROSITE: PS00301; EFACOR GTP: 1.
 CC KM protein biosynthesis; Antibiotic resistance; GTP-binding.
 CC FT NP_BIND 11 18 GTP (BY SIMILARITY).
 CC FT NP_BIND 75 79 GTP (BY SIMILARITY).
 CC FT NP_BIND 129 132 GTP (BY SIMILARITY).
 CC SQ SEQUENCE 652 AA; 72722 MW; EFESL113D562113 CRC64;
 Query Match 40.1%; Score 46.5; DB 1; Length 652;
 Best Local Similarity 55.6%; Pred. No. 13;
 Matches 10; Conservative 1; Mismatches 6; Indels 1; Gaps 1;

OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kedar D., Dumanaki J.P.;
 RT "Cloning of the human and mouse growth factor receptor binding protein
 RT like genes.";
 RT Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Bourette R.P., Arnaud S., Myles G.M., Rohsneider L.R.,
 RA Blanchet J.P., Mouchiroud G.;
 RT "A novel hematopoietic-specific adaptor interacting with the
 RT macrophage-colony-stimulating factor receptor, is implicated in
 RT monocyte/macrophage development.";
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99087328; PubMed=9872323;
 RA Liu S.K., McGlade C.J.;
 RT "Gads is a novel SH2 and SH3 domain-containing adaptor protein that
 RT binds to tyrosine-phosphorylated Shc.";
 RL Oncogene 17:3073-3082(1998).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Iaw C.-L., Ewings M.K., Chaudhary P.M., Solow S.A., Yun T.J.,
 RA Marshall A.J., Hood L., Clark E.A.;
 RT "GrpL, a Grb2-related adaptor protein, interacts with SLP-76 to
 RT regulate NF-AT activation.";
 RN Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Ellis J.N., Ashman C., Burden M., Kilpatrick K.E., Morse M.A.,
 RA Handlin P.A.;
 RT "Grb2, a novel Grb2-related adaptor protein which interacts with the
 RT activated T cell co-stimulatory receptor CD28.";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: INTERACTS WITH SLP-76 TO REGULATE NF-AT ACTIVATION.
 CC BINDS TO TYROSINE-PHOSPHORYLATED SHC.
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE GRB2 / SEM-5 / DRK FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AJ011735; CA009756.1; -;
 CC DR EMBL: AF055465; AAD08803.1; -;
 CC DR EMBL: AF053405; AAC98669.1; -;
 CC DR EMBL: AF129477; AAD41783.1; -;
 CC DR EMBL: AF236118; AAF60318.1; -;
 CC DR MGP: MGI:1333842; Mona.
 CC DR HSSP: P29354; IGR1.
 CC DR InterPro: IPR000980; -;
 CC DR InterPro: IPR001452; -;
 CC DR Pfam: PF00017; SH2; 1.
 CC DR Pfam: PF00018; SH3; 2.
 CC DR PRINTS: PR00452; SH3DOMAIN.
 CC DR PROSITE: PS50001; SH2; 1.
 CC DR PROSITE: PS50002; SH3; 2.
 CC SH2 domain; SH3 domain.
 CC FT DOMAIN 1 56 SH3.
 CC FT DOMAIN 58 149 SH2.
 CC FT DOMAIN 263 322 SH3.
 CC SQ SEQUENCE 322 AA; 36810 MW; 736311D0640CD3D0 CRC64;

Query Match 39.7%; Score 46; DB 1; Length 322;
 Best Local Similarity 50.0%; Pred. No. 8;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 5 NPAMTGRGIRPVGRF 20
 11:1111 :11

DB 297 NPSMTGRLNKLGLF 312

RESULT 11
 ID GRP2_HUMAN STANDARD; PRT; 330 AA.
 AC 057591; 043726;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE GRB2-RELATED ADAPTOR PROTEIN 2 (GADS PROTEIN) (GROWTH FACTOR RECEPTOR
 DE BINDING PROTEIN) (GRB2) (GRB40 ADAPTOR PROTEIN) (GRB-2-LIKE PROTEIN)
 DE (GRB2L) (GRB2) (P38) (HEMATOPOIETIC CELL-ASSOCIATED ADAPTOR PROTEIN
 DE GRP2) (ADAPTOR PROTEIN GRB2).
 GN GRP2 OR GADS OR GRB2L OR GRB2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN 1;
 RC SEQUENCE FROM N.A.
 RP Tissue-Blood;
 RA Freason J.;
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 RN 12;
 RP SEQUENCE FROM N.A.
 RA Kedra D., Dumanski J.P.;
 RT "Cloning of the human and mouse growth factor receptor binding protein
 RT like genes.";
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 RN 13;
 RP SEQUENCE FROM N.A.
 RA Asada H., Ishi N., Sasaki Y., Endo K., Kasai H., Tanaka N.,
 RA Takeshita T., Tsuchiya S., Kono T., Sugamura K.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 RN 14;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99097254; PubMed=9878555;
 RA Qiu M., Hua S., Agrawal M., Li G., Cai J., Chan E., Zhou H., Luo Y.,
 RA Liu M.;
 RT "Molecular cloning and expression of human grb2-2, a novel leukocyte-
 RT specific SH2- and SH3-containing adaptor-like protein that binds to
 RT gab-1.";
 RL Biochem. Biophys. Res. Commun. 253:443-447(1998).
 RN 15;
 RP SEQUENCE FROM N.A.
 RA Teramoto T., Nagashima M., Terai S., Thorgerisson S.S.;
 RT "Grb2, new recruited signaling gene having homology with Grb2.";
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 RN 16;
 RP SEQUENCE FROM N.A.
 RA Law C.-I., Ewings M.K., Chaudhary P.M., Solow S.A., Yun T.J.,
 RA Marshall A.J., Hood L., Clark E.A.;
 RT "Grb2, a Grb2-related adaptor protein, interacts with SLP-76 to
 RT regulate NF-AT activation.";
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RN 17;
 RP SEQUENCE FROM N.A.
 RA Ellis J.H., Ashman C., Burden M., Kilpatrick K.E., Morse M.A.,
 RA Hamblin P.A.;
 RT "GRB2, a novel Grb2-related adaptor protein which interacts with the
 RT activated T cell co-stimulatory receptor.";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN 18;
 RP SEQUENCE FROM N.A.
 RA Burgess J., Whiteley M.;

RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: INTERACTS WITH SLP-76 TO REGULATE NF-AT ACTIVATION.
 CC BINDS TO TYROSINE-PHOSPHORYLATED SHC.
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 SH3 DOMAINS.
 CC -1- SIMILARITY: BELONGS TO THE GRB2 / SEM-5 / DRK FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed, usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: Y18951; CAA7021.1; -;
 DR EMBL: AJ011736; CAA09757.1; -;
 DR EMBL: AF042380; AAC69273.1; -;
 DR EMBL: AF102694; AAD04926.1; -;
 DR EMBL: AF090456; AAD13027.1; -;
 DR EMBL: AF129476; AAD11782.1; -;
 DR EMBL: AF236120; AAF60320.1; -;
 DR EMBL: Z82206; CAB05103.1; ALT_SEQ.
 DR HSSP: P29354; IGRI.
 DR MIM: 604518; -;
 DR InterPro: IPR000980; -;
 DR InterPro: IPR001452; -;
 DR Pfam: PF00017; SH2; 1.
 DR Pfam: PF00018; SH3; 2.
 DR PRINTS: PR00452; SH3DOMAIN.
 DR PROSITE: PS50001; SH2; 1.
 DR PROSITE: PS50002; SH3; 2.
 KW SH2 domain; SH3 domain.
 FT DOMAIN 1 56 SH3.
 FT DOMAIN 58 149 SH2.
 FT DOMAIN 271 330 SH3.
 SQ SEQUENCE 330 AA; 37909 MW; 74F4C8D0EB56D55 CRC64;

Query Match 39.7%; Score 46; DB 1; Length 330;
 Best Local Similarity 50.0%; Pred. No. 8.2;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 5 NPAMTGRGIRPVGRF 20
 11:1111 :11

DB 305 NPSMTGRLNKLGLF 320

RESULT 12
 ID UVR4_HAETN STANDARD; PRT; 943 AA.
 AC P44410; 048151;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE EXCINUCLEASE ABC SUBUNIT A.
 GN UVR4 OR H10249.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Haemophilus.
 OX NCBI_TaxID=727;
 RN 1;
 RP SEQUENCE FROM N.A.
 RC STRAIN-RD / KW20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McInerney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-T., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Heddlom E., Cotton M.D.,
 RA Uiterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.;

RA Venter J.C.;
 RT "RpoIe-genome random sequencing and assembly of Haemophilus influenzae
 RL Science 269:496-512(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NTHI N187;
 RA MEDLINE=97080495; PubMed=89218440;
 RA de la Mora M.L., Hendrixson D.R., St Gene J.W. III;
 RT "Isolation and characterization of the Haemophilus influenzae uvra
 RL gene.";
 RN Gene 177:23-28(1996).
 RL [3]
 RP SEQUENCE OF 1-71 FROM N.A.
 RC STRAIN-NTHI TN106;
 RA MEDLINE=94341556; PubMed=8063092;
 RA Jarosik G.P., Hansen E.J.;
 RT "Cloning and sequencing of the Haemophilus influenzae ssb gene
 RL encoding single-strand DNA-binding protein.";
 RL Gene 146:101-103(1994).
 CC -1- FUNCTION: THE ABC EXCISION NUCLEASE IS A DNA REPAIR ENZYME THAT
 CC CATALYZES THE EXCISION REACTION OF UV-DAMAGED NUCLEOTIDE SEGMENTS
 CC PRODUCING OLIGOMERS HAVING THE MODIFIED BASE(S). UVRA IS AN ATPASE
 CC AND A DNA-BINDING PROTEIN THAT PREFERENTIALLY BINDS SINGLE-
 CC STRANDED OR UV-IRRADIATED DOUBLE-STRANDED DNA (BY SIMILARITY).
 CC -1- SUBUNIT: CONSISTS OF THREE SUBUNITS; UVRA, UVRA AND UVRC.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
 CC (ABC TRANSPORTERS). CONTAINS TWO ABC DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U32711; AAC21915.1; -;
 CC EMBL: U33877; AAC44592.1; -;
 CC EMBL: U04897; AAA60462.1; -;
 DR TIGR: H10249; -;
 DR InterPro: IPR001617; -;
 DR Pfam: PF00005; ABC_tran. 2;
 DR PROSITE: PS00211; ABC_TRANSPORTER. 2;
 KM SOS response: Excision nuclease; DNA repair; ATP-binding; Repeat;
 KM DNA-binding; Zinc-finger.
 FT NP_BIND 31 38 ATP (POTENTIAL).
 FT NP_BIND 640 647 ATP (POTENTIAL).
 FT ZN_FING 253 280 C4-TYPE.
 FT ZN_FING 740 766 C4-TYPE.
 FT CONFLICT 163 163 V -> L (IN REF. 2).
 FT CONFLICT 236 236 E -> D (IN REF. 2).
 FT CONFLICT 425 425 R -> K (IN REF. 2).
 FT CONFLICT 463 463 I -> M (IN REF. 2).
 FT CONFLICT 514 514 E -> Q (IN REF. 2).
 FT CONFLICT 661 661 A -> T (IN REF. 2).
 FT CONFLICT 928 928 T -> E (IN REF. 2).
 FT CONFLICT 935 943 FLAKILKRP -> PLP (IN REF. 2).
 SO SEQUENCE 943 AA; 104366 MW; 4DBA0DCFA602D65 CRC64;

Query Match 39.7%; Score 46; DB 1; Length 943;
 Best Local Similarity 37.8%; Pred. No. 23;
 Matches 14; Conservative 0; Mismatches 5; Indels 18; Gaps 2;
 QY 1 TPDINPAWYTG-----RGIRPVGRF 20
 DB 696 TPRSNPATYTGTFPIRELFAGVPEARARGYNP-GRF 731
 RESULT 13
 R167_YEAST

ID R167_YEAST STANDARD; PRT; 482 AA.
 AC P39743;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE REDUCED VIABILITY UPON STARVATION PROTEIN 167.
 GN RVS167 OR YDR368W OR D9509.8.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-X2180;
 RX MEDLINE=93330299; PubMed=8336735;
 RA Bauer F., Undack M., Aigle M., Cronzet N.;
 RT "Alteration of a yeast SH3 protein leads to conditional viability
 RL with defects in cytoskeletal and budding patterns.";
 RL Mol. Cell. Biol. 13:5070-5084(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Dietrich F.S., Mulligan J., Allen E., Araujo R., Aviles E.,
 RA Berno A., Carpenter J., Chen E., Cherry J.M., Chung E., Duncan M.,
 RA Hunkeler-Smith S., Hymen R., Komp C., Lashkari D., Lew H., Lin D.,
 RA Mosedale D., Nakahara K., Namath A., Oelner P., Oh C., Petel F.K.,
 RA Roberts D., Schramm S., Schroeder M., Shogren T., Shroff N.,
 RA Viant A., Yelton M., Botstein D., Davis R.R.;
 RL Submitted (JUL-1995) to the EMBL/GenBank/DDay databases.
 RN [3]
 RP ACTIN-BINDING.
 RX MEDLINE=95236199; PubMed=7719850;
 RA Amberg D.C., Basart E., Botstein D.;
 RT "Defining protein interactions with yeast actin in vivo.";
 RL Nat. Struct. Biol. 2:28-35(1995).
 CC -1- FUNCTION: COMPONENT OF A CYTOSKELETAL STRUCTURE THAT IS REQUIRED
 CC FOR THE FORMATION OF ENDOCYTIC VESICLES AT THE PLASMA MEMBRANE
 CC LEVEL. COULD BE IMPLICATED IN CYTOSKELETAL REORGANIZATION IN
 CC RESPONSE TO ENVIRONMENTAL STRESSES AND COULD ACT IN THE BUDDING
 CC SITE SELECTION MECHANISM. BINDS TO ACTIN.
 CC -1- SIMILARITY: TO YEAST RVS167.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M92092; AAA35051.1; -;
 CC EMBL: U32274; AAB64830.1; -;
 DR PIR: S40887; S40887.
 DR HSSP: P04002; IATF.
 DR SGD: S0002796; RVS167.
 DR InterPro: IPR001452; -;
 DR Pfam: PF00018; SH3. 1.
 DR PRINTS: PR00452; SH3DOMAIN.
 DR PROSITE: PS50002; SH3. 1.
 KM Cytoskeleton; SH3 domain; Transmembrane; Actin-binding.
 FT DOMAIN 292 427 ALA/GLY/PRO-RICH.
 FT TRANSMEM 344 367 POTENTIAL.
 FT DOMAIN 421 482 SH3.
 SO SQUENCE 482 AA; 52774 MW; 3F0AB53EBC095A5B CRC64;

Query Match 38.8%; Score 45; DB 1; Length 482;
 Best Local Similarity 47.6%; Pred. No. 17;
 Matches 10; Conservative 3; Mismatches 2; Indels 6; Gaps 2;
 QY 1 TPDINPAWYTG-----GIRP 16
 DB 454 TPDVN-EMWTGRYNGGQGVPP 473

```

RESULT 14
EXL1_HUMAN STANDARD; PRT; 676 AA.
ID EXL1_HUMAN
AC Q92935;
DR 15-JUL-1998 (Rel. 36, Created)
DR 15-JUL-1998 (Rel. 36, Last sequence update)
DR 01-OCT-2000 (Rel. 40, Last annotation update)
DE EXOSTOSIN-LIKE 1 (EXOSTOSIN-L) (MULTIPLE EXOSTOSIS-LIKE PROTEIN).
GN EXL1 OR EXTL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97189339; PubMed=9037597;
RA Wise C.A., Clines G.A., Massa H., Trask B.J., Lovett M.;
RT "Identification and localization of the gene for EXTL, a third member
RT of the multiple exostoses gene family."
RL Genome Res. 7:10-16(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Xu L., Deng H.X., Xia J.H., Pan Q., Liu C.Y.;
RT "Mutations of the EXT genes in hereditary multiple exostoses in
RT Chinese."
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Wuys W., Spleker N., Van Roy N., De Paepe A., De Boulle K.,
RA Willems P.J., Van Hul W., Versteeg R., Speleman F.;
RT "Refined physical mapping and genomic structure of the EXTL1 gene."
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: APPEARS TO BE A TUMOR SUPPRESSOR (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ENDOPLASMIC
CC RETICULUM (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE EXOSTOSIN FAMILY.
CC -----
CC This SWISS-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: U67191; AAC51141.1; -.
DR EMBL: AF083633; AAD02840.1; -.
DR EMBL: AF083623; AAD02840.1; JOINED.
DR EMBL: AF083624; AAD02840.1; JOINED.
DR EMBL: AF083625; AAD02840.1; JOINED.
DR EMBL: AF083626; AAD02840.1; JOINED.
DR EMBL: AF083627; AAD02840.1; JOINED.
DR EMBL: AF083628; AAD02840.1; JOINED.
DR EMBL: AF083629; AAD02840.1; JOINED.
DR EMBL: AF083630; AAD02840.1; JOINED.
DR EMBL: AF083631; AAD02840.1; JOINED.
DR EMBL: AF083632; AAD02840.1; JOINED.
DR EMBL: AF153980; AAF73172.1; -.
DR EMBL: AF153991; AAF73172.1; JOINED.
DR MIM: 601738; -.
KW Anti-oncogene; Multigene family; Transmembrane; Signal-anchor.
FT TRANSMEM 10 30
FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT CARBOHYD 269 269
FT N-LINKED (GLCNAC... ) (POTENTIAL).
SO SEQUENCE 676 AA; 74673 MW; B5E006A8762E5633 CRC64;

```

Query Match 38.8%; Score 45; DB 1; Length 676;
 Best Local Similarity 45.0%; Pred. No. 23;
 Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 TPDIINPARYTGRIPIVGRF 20
 : | | | | | | | | | |
 Db 400 SPDPFFRYLQGSRSRGRF 419

```

RESULT 15
DPOOL_THEGO STANDARD; PRT; 773 AA.
ID DPOOL_THEGO
AC P56689;
DR 15-JUL-1999 (Rel. 38, Created)
DR 15-JUL-1999 (Rel. 38, Last sequence update)
DR 15-JUL-1999 (Rel. 38, Last annotation update)
DE DNA POLYMERASE (EC 2.7.7.7) (TO POL).
GN POL OR POLA.
OS Thermococcus gorgonarius.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Thermococcus.
OX NCBI_TaxID=71997;
RN [1]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX MEDLINE=9919230; PubMed=10097083;
RA Hopfinger K.-P., Eichinger A., Engh R.A., Laue F., Ankenbauer W.,
RA Huber R., Angerer B.;
RT "Crystal structure of a thermostable type B DNA polymerase from
RT Thermococcus gorgonarius."
RL Proc. Natl. Acad. Sci. U.S.A. 96:3600-3605(1999).
CC -1- FUNCTION: IN ADDITION TO POLYMERASE ACTIVITY, THIS DNA POLYMERASE
CC EXHIBITS 3' TO 5' EXONUCLEASE ACTIVITY.
CC -1- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE -
CC N PYROPHOSPHATE + DNA(N).
CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
CC PDB: 1TGO; 22-MAR-99.
DR InterPro: IPR002064; -.
DR Pfam: PF00136; DNA_pol_B. 1.
DR PRINTS: PR00106; DNAPOLB.
DR PROSITE: PS00116; DNA_POLYMERASE_B. 1.
KW transferase; DNA-directed DNA polymerase; DNA replication;
KW DNA-binding; Hydroxylase; Nuclease; Exonuclease; Multifunctional enzyme;
KW 3D-structure.
FT DISUFID 428 442
FT DISUFID 506 509
SO SEQUENCE 773 AA; 89812 MW; F67AF04E875FBE44 CRC64;

```

Query Match 38.8%; Score 45; DB 1; Length 773;
 Best Local Similarity 44.4%; Pred. No. 26;
 Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Search completed: April 17, 2001, 15:48:49
 Job time: 537 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: April 17, 2001, 15:48:10 ; Search time 115.07 Seconds
(without alignments)
20.372 Million cell updates/sec

Title: US-09-446-543a-64

Perfect score: 113

Sequence: 1 TPDINPAMYASRGIRPVGRF 20

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP_TREMBL_15:
2: SP_archaea:
3: SP_bacteria:
4: SP_fungi:
5: SP_human:
6: SP_invertebrate:
7: SP_mhc:
8: SP_mhc1:
9: SP_mhc2:
10: SP_mhc3:
11: SP_mhc4:
12: SP_mhc5:
13: SP_mhc6:
14: SP_virus:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	84	74.3	117	13	Q9W624
2	54	47.8	790	10	Q9M371
3	51	45.1	54	4	Q9U9P9
4	51	45.1	465	4	Q60687
5	47.5	42.0	333	2	Q9KJ10
6	47	41.6	284	1	Q50128
7	47	41.6	333	2	Q9PH76
8	47	41.6	938	10	Q9ZVE3
9	46.5	41.2	779	3	Q9P5J9
10	46	40.7	1296	5	Q22452
11	45.5	40.3	1501	10	Q9SD86
12	45	39.8	153	9	Q9T133
13	45	39.8	207	5	Q27455
14	45	39.8	390	2	Q9PH18
15	45	39.8	419	4	Q9Y276
16	45	39.8	430	1	Q27142
17	45	39.8	540	10	Q9L620
18	45	39.8	767	5	Q20170
19	44.5	39.4	259	2	Q9L1R4

20	44	38.9	276	2	Q53929
21	44	38.9	284	1	Q9V003
22	44	38.9	302	2	P96274
23	44	38.9	309	5	Q17234
24	44	38.9	428	2	Q9PEK7
25	44	38.9	527	5	Q76383
26	43.5	38.5	374	2	Q53319
27	43	38.1	402	2	Q9K1W3
28	43	38.1	466	2	Q54042
29	43	38.1	743	10	Q9SKV6
30	43	38.1	1359	10	Q9SW46
31	43	38.1	1364	10	Q9SW45
32	42.5	37.6	345	2	Q9K8V2
33	42.5	37.6	1540	10	Q3SD90
34	42.5	37.6	3944	5	Q18667
35	42	37.2	72	8	Q9P6D9
36	42	37.2	128	2	P74747
37	42	37.2	132	2	Q9P6J6
38	42	37.2	184	2	Q86838
39	42	37.2	216	2	Q9P8U7
40	42	37.2	219	10	Q9SL72
41	42	37.2	221	2	Q9P8S3
42	42	37.2	232	2	Q9R221
43	42	37.2	343	3	Q74569
44	42	37.2	459	3	Q94284
45	42	37.2	512	3	Q94269

ALIGNMENTS

RESULT 1
ID Q9W624 PRELIMINARY; PRT; 117 AA.
AC Q9W624;
DT 01-NOV-1999 (TREMURel. 12, Created)
DM 01-NOV-1999 (TREMURel. 12, Last sequence update)
DN 01-NOV-1999 (TREMURel. 12, Last annotation update)
DE C-RF AMTIDE PRECURSOR.
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Cyprininae; Carassius.
OX NCBI_TaxID=7957;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Satake H., Minakata H., Fujimoto M.;
RT "Carassius Rhamde (C-RF amide).";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB020024; BAA76662.1; D5DC4CB22038C2B0 CRC64;
SO SEQUENCE 117 AA; 12879 MW; D5DC4CB22038C2B0 CRC64;

Query Match 74.3%; Score 84; DB 13; Length 117;
Best Local Similarity 60.0%; Pred. No. 1.2e-05;
Matches 12; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
QY 1 TPDINPAMYASRGIRPVGRF 20
Db 56 SPIDPFWYVGRVGRVGRF 75
RESULT 2
ID Q9M371 PRELIMINARY; PRT; 790 AA.
AC Q9M371;
DT 01-OCT-2000 (TREMURel. 15, Created)
DM 01-OCT-2000 (TREMURel. 15, Last sequence update)
DN 01-OCT-2000 (TREMURel. 15, Last annotation update)
DE HYPOTHETICAL 87.4 KDA PROTEIN.
GN F15G16.60.
OS Arabidopsis thaliana (Mouse-ear cress).

```

OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
ON NCBI_TaxID=3702;
RN
  11
  RP SEQUENCE FROM N.A.
  RA De Haan M., Maaree A.C., Grivell L.A., Mewes H.W., Lemcke K.,
  RA Mayer K.F.X., Quetier F., Salanoubat M.;
  RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
  RN
    [2]
  RP SEQUENCE FROM N.A.
  RA EU Arabidopsis sequencing project;
  RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
  DR EMBL: AL132959; CAB71097.1; -.
  KM Hypothetical protein.
  SQ SEQUENCE 790 AA; 87375 MW; B222724B75690F30 CRC64;

Qy
  1 1111111111
  Db 366 PPHNPRYGSRGLOPHGRV 384

RESULT 3
O9UUF9 PRELIMINARY; PRT; 54 AA.
ID AC O9UUF9;
DT 01-MAY-2000 (TREMBLrel. I3, Created)
DT 01-MAY-2000 (TREMBLrel. I3, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. I3, Last annotation update)
DE D1479J7.3 (SUSHI-REPEAT PROTEIN (SRPUL)) (FRAGMENT).
GN D1479J7.3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
ON
  11
  RP SEQUENCE FROM N.A.
  RA Lawlor S.;
  RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
  DR EMBL: AL035608; CAB55682.1; -.
  FT NON_TER 54
  SQ SEQUENCE 54 AA; 6110 MW; E2F3C39F7B961A9F CRC64;

Query Match 45.1%; Score 51; DB 4; Length 54;
Best Local Similarity 50.0%; Pred. No. 0.64;
Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Oy
  1 1111111111
  Db 18 TPAVPTWYAGSGYTP 33

RESULT 4
O60687 PRELIMINARY; PRT; 465 AA.
ID AC O60687;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE SUSHI-REPEAT PROTEIN.
GN SRPUL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
ON
  11
  RP SEQUENCE FROM N.A.

```

```

RA Kurosawa H., Inukati T., Imaba T., Goli K., Chang K.-S., Sinjo T.,
RA Rakeshaw K.M., Neeve C.W., Lock T.A.;
DR Submitted (APR-1998) to the EMBL/Genbank/DBJ databases.
RL EMBL; AF060567; AAC15765.1; -.
DR INTERPRO: IPR000436; -.
DR INTERPRO: IPR001128; -.
DR PFM: PFO0084; Sushl; 3.
DR PROSITE: PS00086; CYTOCHROME_P450; UNKNOWN_1.
SO SEQUENCE 465 AA; 52971 MW; 4D752B187F3EFPB8 CRC64;

Query Match          45.1%; Score 51; DB 4; Length 465;
Best Local Similarity 50.0%; Pred. No. 6.4;
Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

OY      1 TPIDINPAWVSRGIRP 16
         || : | ||| | |
DB       18 TPATPTWVGSGGY 33

RESULT    5
OGRJTO   OGRJTO           PRELIMINARY;     PRT;      333 AA.
AC      OGRJTO;
AD      OGRJTO;
DT      01-MAY-2000 (TREMBLrel. 13, Created)
DT      01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT      01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE      HYPOTHEICAL 36.3 KDA PROTEIN.
GN      SCF73.06C.
OS      Streptomyces coelicolor.
OC      Bacteria; Firmicutes; Actinobacteria; Actinomycetales;
OC      Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX      MBL_TaxId=1902;
RN      RN
RP      SEQUENCE FROM N.A.
RC      STRAIN-A3(2);
RA      Seeger K.J., Harris D.;
RL      Submitted (OCT-1999) to the EMBL/Genbank/DBJ databases.
[2]
RN      RN
RP      SEQUENCE FROM N.A.
RC      STRAIN-A3(2);
RA      Parkhill J., Bartell B.G., Rajandream M.A.;
RL      Submitted (OCT-1999) to the EMBL/Genbank/DBJ databases.
[3]
RN      RN
RP      SEQUENCE FROM N.A.
RC      STRAIN-A3(2);
RA      Redenbach M., Kieser H.M., Denaplatte D., Eichner A., Cullum J.,
RA      Kinashi H., Hopwood D.A.;
RA      "A set of ordered cosmids and a detailed genetic and physical map for
RI      the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL      Mol. Microbiol. 21:77-96(1996).
DR      EMBL; ALI21746; CAB57411.1; -.
KW      Hypothetical protein.
SQ      SEQUENCE 333 AA; 36312 MW; 884FF2861837FD76 CRC64;

Query Match          42.0%; Score 47.5; DB 2; Length 333;
Best Local Similarity 35.5%; Pred. No. 16;
Matches 11; Conservative 3; Mismatches 6; Indels 11; Gaps 1;

OY      1 TPIDINPAW-----YASRGIRPVORF 20
         || |||              | : | || |||
DB       19 TPMEEPARTEAIGTWIEARLANAGILPIGRW 49

RESULT    6
O5O128   O5O128           PRELIMINARY;     PRT;      284 AA.
AC      O5O128;
DT      01-JUN-1998 (TREMBLrel. 06, Created)
DT      01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT      01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE      HYPOTHEICAL 32.3 KDA PROTEIN PH120.
```

OS PH1420.
 OS Pyrococcus horikoshii.
 OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
 NC NCBI_TaxID=53953;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OT3;
 RA MEDLINE=98344137; PubMed=9679194;
 RA Kawarabayashi Y., Sawada M., Horikawa H., Halkawa Y., Hino Y.,
 RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Hagi Y.,
 RA Sakai M., Ogura K., Otsuka R., Hakezawa H., Takamiya M., Ohfuku Y.,
 RA Funahashi T., Tanaka T., Kudo Y., Yamazaki Y., Kishida H., Oguchi A.,
 RA Aoki K.-I., Yoshizawa T., Nakamura Y., Kobb F.T., Horikoshi K.,
 RA Masuchi Y., Shizuya H., Kikuchi H.;
 RT Complete sequence and gene organization of the genome of a hyper-
 RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3.".
 RL DNA Res. 5:55-76(1998).
 DR EMBL: AP000006; BAA30526.1; -.
 DR INTERPRO: IPR001066; -.
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; UNKHOWM_1.
 KW Hypothetical protein.
 SQ SEQUENCE 284 AA; 32319 MW; 8E0E7BC3711D3815 CRC64;

Query Match 41.6%; Score 47; DB 1; Length 284;
 Best Local Similarity 50.0%; Pred. No. 16;
 Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 PDHPAWYASRGIRPYGR 19
 DB 217 PYLEPTFYALRGLELLGR 234

RESULT 7
 Q9PH76 PRELIMINARY; PRT: 333 AA.
 AC Q9PH76;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, last annotation update)
 DE HYDROXYBENZONATE OCTAPRENT/TRANSFERASE.
 GH XE0068.
 OS Xylella fastidiosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
 OC Xylella.
 NC NCBI_TaxID=23711;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=9A5C;
 RA MEDLINE=20365717; PubMed=10910347;
 RA Simpson A.J.G., Reinach F.C., Artuda P., Abreu F.A., Acencio M.,
 RA Alvaranga R., Alves L.M.C., Araya J.E., Baiz G.S., Baptista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carter H.,
 RA Coutinho N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Helo C.M.,
 RA Fagundes A.P., Ferreira A.J.S., Ferreira V.C.A., Ferriz J.A.,
 RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
 RA Garlier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 RA Ho P.L., Hohnselt J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
 RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 RA Machado M.A., Madela A.M.B.H., Madela H.M.F., Martino C.L.,
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
 RA Moon D.H., Nagai M.B., Nascimento A.L.T.O., Netto L.E.S.,
 RA Nhani A.J., Nobrega F.G., Nunes L.R., Oliveira M.A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
 RA Peixoto B.R., Pereira G.A.G., Pereira R.A.Jr., Pesquero J.B.,
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
 RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,

RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako H.H.,
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Zeldanis J., Zetel J.C.;
 RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
 RL Nature 406:151-157(2000).
 DR EMBL: AE003860; AAF82881.1; -.
 DR INTERPRO: IPR000537; -.
 DR PFAM: PF01040; UBLA; 1.
 DR PROSITE: PS00943; UBLA; UNKNOWN_1.
 SQ SEQUENCE 333 AA; 37931 MW; ECF3F4716C962895 CRC64;

Query Match 41.6%; Score 47; DB 2; Length 333;
 Best Local Similarity 53.3%; Pred. No. 19;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 4 INPAMWASRGIRPYG 18
 DB 54 LDPYWKLRGDRPYG 68

RESULT 8
 Q9ZVE3 PRELIMINARY; PRT: 938 AA.
 AC Q9ZVE3;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, last sequence update)
 DT 01-MAY-1999 (TREMBLrel. 10, last annotation update)
 DE 19F8.5 PROTEIN.
 GH 19F8.5.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 NC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV, COLUMBIA;
 RA Rounsley S.D., Lin X., Kaul S., Shen T.P., Fujii C.Y., Mason T.M.,
 RA Shen M., Roeding C.M., Fraser C.M., Somerville C.R., Venter J.C.;
 RT Arabidopsis thaliana chromosome II BAC 19F8 genomic sequence.";
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC005561; AAC98465.1; -.
 SQ SEQUENCE 938 AA; 105542 MW; 2A42C66E9C590B78 CRC64;

Query Match 41.6%; Score 47; DB 10; Length 938;
 Best Local Similarity 66.7%; Pred. No. 56;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 5 NPAMWASRGIRPYGR 19
 DB 60 HPAMWASRGIRPYGR 74

RESULT 9
 Q9P5J9 PRELIMINARY; PRT: 779 AA.
 AC Q9P5J9;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, last annotation update)
 DE RELATED TO BCS1 PROTEIN PRECURSOR.
 GN B23121.300.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Sordariales; Sordariaceae; Neurospora.
 NC NCBI_TaxID=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Schulte U., Altmann J., Brandt P., Fatmann B., Holland R.,
 RA Harknuta G., Mewes H.W., Mannhaupt G.;
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]

RP SEQUENCE FROM N.A.
 RA German Neurospora genome project;
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL356172; CAB91698.1; -
 SQ SEQUENCE 779 AA; 85796 MW; CA7891402DFEBE30 CRC64;

Query Match 41.2%; Score 46.5; DB 3; Length 779;
 Best Local Similarity 58.8%; Pred. No. 55;
 Matches 10; Conservative 2; Mismatches 2; Indels 3; Gaps 1;

QY 1 TPDPINPA---WYASRG1 14
 1 :||| |||:|
 DB 286 TDYINPATRRYANRGI 302

RESULT 10

ID Q22452 PRELIMINARY; PRT; 1296 AA.

AC Q22452;

DT 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

DE SIMILAR TO AGRIN AND FOLISTATIN.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Rhabditidae; Pelodierinae; Caenorhabditis.

NCBI_Taxid=6239;

RP SEQUENCE FROM N.A.

RA MEDLINE-94150718; PubMed-7906398;

RA Willson R., Ainscough R., Anderson K., Baynes C., Berks M.,

RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,

RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,

RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,

RA Jones M., Kershaw J., Kristen J., Laister N., Latreille P.,

RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,

RA Parsons J., Percy C., Rikken L., Roopra A., Saunders D., Shownkeen R.,

RA Smailson N., Smith A., Sonhammer E., Staden R., Sulston J.,

RA Thayer-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,

RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.,

RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.

RT elegans".

RL Nature 368:32-38(1994).

RN [2]

RP SEQUENCE FROM N.A.

RA Du Z.;

RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RA Waterston R.;

RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.

DR HSSP: P37109; IPCE.

DR INTERPRO: IPR000561; -

DR INTERPRO: IPR001239; -

DR INTERPRO: IPR001791; -

DR INTERPRO: IPR002049; -

DR INTERPRO: IPR002350; -

DR PFAM: PF00050; Kazal; 9.

DR PFAM: PF00053; laminin_EGF; 2.

DR PFAM: PF00054; laminin_G; 1.

DR PRINTS: PR00280; KAZALINHR.

DR PROSITE: PS00022; EGF_1; UNKNOWN_2.

DR PROSITE: PS01186; EGF_2; UNKNOWN_1.

DR PROSITE: PS01248; LAMININ_TYPE_EGF; 1.

SQ SEQUENCE 1296 AA; 145178 MW; 05094BC185839690 CRC64;

QY 1 TPDPINPAWYASRGIRPVGRF 20
 1 :||| ||| :||
 DB 975 TPDPINSDWYFSR--KDINRF 992

RESULT 11

ID Q9SD86 PRELIMINARY; PRT; 1501 AA.

AC Q9SD86;

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)

DE HYPOTHETICAL 164.4 KDA PROTEIN.

GN F13624.180.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;

OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;

OC Brassicales; Brassicaceae; Arabidopsids.

NCBI_Taxid=3702;

RP SEQUENCE FROM N.A.

RA Bevan M., Van Der Schueren J., Chuang Y.J., Voet M., Robben J.,

RA Volckaert G., Bancroft I., Mewes H.W., Lemcke K., Mayer K.F.X.;

RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA Arabidopsis sequencing project;

RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: AL133421; CAB62610.1; -

DR HYPOTHETICAL protein.

SQ SEQUENCE 1501 AA; 164360 MW; EAD2AEDC9EBE5D2 CRC64;

Query Match 40.3%; Score 45.5; DB 10; Length 1501;
 Best Local Similarity 34.5%; Pred. No. 1.6e+02;
 Matches 10; Conservative 6; Mismatches 4; Indels 9; Gaps 2;

QY 1 TPDPINPAWYASRG-----IRPV---GRF 20
 1 :||| ||| :||
 DB 1184 SPDMAPSWISQVTFPKNGLVQVPVNDGRF 1212

RESULT 12

ID Q9T133 PRELIMINARY; PRT; 153 AA.

AC Q9T133;

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)

DE ENDONUCLEASE.

GN 3.

OS Bacteriophage phi-Yeo3-12.

OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Podoviridae;

OC T7-like phages.

NCBI_Taxid=110457;

RN [1]

RP SEQUENCE FROM N.A.

RA Pajunen M.I., Kiljunen S.J., Skurnik M.;

RT "Complete genomic sequence of the lytic bacteriophage phi-Yeo3-12 of

RT versinia enterocolitica serotype 0:3";

RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: AJ251805; CAB63604.1; -

KW Endonuclease.

SQ SEQUENCE 153 AA; 17640 MW; 211571BBD86C641D CRC64;

Query Match 39.8%; Score 45; DB 9; Length 153;
 Best Local Similarity 66.7%; Pred. No. 17;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 9 YASRGIRPVGRF 20
 1 :||| ||| :||

DB 5 YAAAGVRYGAF 16

RESULT 13

ID Q27455 PRELIMINARY; PRT: 207 AA.

AC Q27455: 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)

DE PHOSPHOENOLPYRUVATE CARBOXYKINASE (EC 4.1.1.32)

DE (PHOSPHOENOLPYRUVATE CARBOXYKINASE (GTP))

DE (PHOSPHOENOLPYRUVATE CARBOXYLASE) (PHOSPHOPYRUVATE CARBOXYLASE)

DE (PEPCK) (FRAGMENT).

OS Brachycentrus nigrosoma.

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Phryganea; Neoptera; Endopterygota; Trichoptera; Integrilipalia;

OC Limnephilidae; Brachycentridae; Brachycentrus.

OX NCBI_TaxID:41036;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE-97036855; PubMed-8882502;

RA Friedlander T.P., Regier J.C., Mitter C., Wagner D.L.;

RT "A nuclear gene for higher level phylogenetics: phosphoenolpyruvate carboxykinase tracks mesozoic-age divergences within Lepidoptera (Insecta)."

RL Mol. Biol. Evol. 13:594-604(1996).

DR EMBL; U28445; AAB40362.1; -.

DR INTERPRO; IPR000364; -.

DR PFAM; PF00821; PEPCK.1.

DR PROSITE; PS00505; PEPCK_GTP.1.

KW Pyruvate; lyase.

FT NON_TER 1

FT NON_TER 207

SQ SEQUENCE 207 AA; 22123 MW; 3488D5E47B563E8 CRC64;

Query Match 39.8%; Score 45; DB 5; Length 207;

Best Local Similarity 37.0%; Pred. No. 23;

Matches 10; Conservative 3; Mismatches 4; Indels 10; Gaps 1;

QY 2 PDINPAMYASRG1-----RPVG 18

DB 146 PTIDPAMESSAGYPSALIFGGRPG 172

RESULT 14

ID Q9PH18 PRELIMINARY; PRT: 390 AA.

AC Q9PH18: 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)

DE CYSTEINE SYNTHASE.

GN XPO128.

OS Xylella fastidiosa.

OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;

OC Xylella.

OX NCBI_TaxID:2371;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-9ASC;

RX MEDLINE-20365717; PubMed-10910347;

RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M., Alvarenga R., Alves L.M.C., Araya J.E., Baia J.S., Baptista C.S., Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Bionis M.R.S., Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carter H., Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa Neto C.M., Coutinho L.B., Cistofani M., Discheto R., Docena C., El-Dorri H., Facchini A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A., Fraga J.S., Frasca S.C., Franco M.C., Frohne M., Furlan L.R., Gandler M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A., Ho P.L., Hohlseil J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,

RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C., Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A., Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Matsukuma A.Y., Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y., Menck C.F.M., Miracca E.C., Miyaki C.Y., Montelero-Vitorello C.B., Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S., Nhami A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A., de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A., Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B., Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M., de Rosa V.B. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E., da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr., da Souza A.P., Terenzi M.F., Trufil D., Tsai S.M., Tsubako M.H., Vallada H., Van Sluys M.A., Verjovsky-Almeida S., Vettore A.L., Zago M.A., Zatz M., Melandis J., Setubal J.C.;

RT "The genome sequence of the plant pathogen Xylella fastidiosa."

RL Nature 406:151-157(2000).

DR EMBL; AE003866; AAF82941.1; -.

DR INTERPRO; IPR001926; -.

DR PFAM; PF00291; PALP.1.

SQ SEQUENCE 390 AA; 42479 MW; D294072E6A55188A CRC64;

Query Match 39.8%; Score 45; DB 2; Length 390;

Best Local Similarity 70.0%; Pred. No. 45;

Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 5 NDPMYASRG1 14

DB 354 SPWYAAHGI 363

RESULT 15

ID Q9Y276 PRELIMINARY; PRT: 419 AA.

AC Q9Y276: 01-NOV-1999 (TREMBLrel. 12, Created)

DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)

DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)

DE H-BCSL.

GN BCSL.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID:9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE-BRAIN;

RX MEDLINE-99097350; PubMed-9878253;

RA Petruzzella V., Tiranti V., Fernandez P., Ianna P., Carrozzo R., Zeylani M.;

RT "Identification and characterization of human cDNAs specific to BCSL, PETL12, SCOL, COX15, and COX11, five genes involved in the formation and function of the mitochondrial respiratory chain."

RL Genomics 54:494-504(1998).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE-BRAIN;

RX MEDLINE-96207227; PubMed-8619474;

RA Andersson B., Wentland M.A., Ricafrente J.Y., Liu W., Gibbs R.A.;

RT "A 'double adaptor' method for improved shotgun library construction."

RL Anal. Biochem. 236:107-113(1996).

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE-BRAIN;

RX MEDLINE-97264341; PubMed-9110174;

RA Yu W., Andersson B., Worley K.C., Munzy D.M., Ding Y., Liu W., Ricafrente J.Y., Wentland M.A., Lennon G., Gibbs R.A.;

RT "Large-scale concatenation cDNA sequencing."

RL Genome Res. 7:353-358(1997).

DR EMBL; AF026849; AAD08638.1; -.

Tue Apr 17 15:46:21 2001

DR EMBL: AF038195; AAB97365.1; -.
DR INTERPRO: IPR001939; -.
DR PFM: PF00004; AAA; 1.
KW Hypothetical protein.
SO SEQUENCE 419 AA; 47534 MW; 7F0F98BA62F2CBB8 CRC64;

Query Match 39.8%; Score 45; DB 4; Length 419;
Best local Similarity 70.0%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 5 NPAMYASRG1 14
|| || || ||
Db 211 NPKMTDRGI 220

Search completed: April 17, 2001, 15:48:11
Job time: 564 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 17, 2001, 15:45:57 ; Search time 70.08 Seconds

(without alignments)
19.613 Million cell updates/sec

Title: US-09-446-543a-64

Sequence: 1 TPDINPAMYASRGIRPVGRF 20

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_67:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	47.8	790	2 T47959	hypothetical prote
2	50.5	44.7	664	2 F83376	conserved hypothet
3	47	41.6	284	2 F71015	hypothetical prote
4	47	41.6	333	2 H82852	hydroxybenzoate oc
5	46.5	41.2	779	2 T49717	related to BCS1 pr
6	46	40.7	1296	2 T16859	hypothetical prote
7	45.5	40.3	1501	2 T45623	hypothetical prote
8	45	39.8	390	2 C82844	cysteine synthase
9	45	39.8	430	1 B69009	conserved hypothet
10	45	39.8	767	2 T21969	hypothetical prote
11	44	38.9	276	2 D70817	hypothetical prote
12	44	38.9	284	2 A75117	hypothetical prote
13	44	38.9	302	2 B70631	hypothetical prote
14	44	38.9	309	2 T32376	hypothetical prote
15	44	38.9	428	2 F81694	hypothetical prote
16	44	38.9	476	2 G64720	pyruvate dehydroge
17	44	38.9	527	2 T33175	probable amino aci
18	44	38.9	719	2 S61046	hypothetical prote
19	43.5	38.5	374	2 G70947	ABPI protein - yea
20	43	38.1	250	2 C83292	hypothetical prote
21	43	38.1	220	2 C83400	probable glutathio
22	43	38.1	376	2 A48197	hypothetical prote
23	43	38.1	376	2 B48197	opsin, ocular - A
24	43	38.1	468	2 C83160	opsin, lateral eye
25	43	38.1	1359	2 T10235	nitrite extrusion
26	43	38.1	1364	2 T10236	xanthine dehydroge
27	42.5	37.6	345	2 D84012	N-acetylglutamate
28	42.5	37.6	1540	2 T45619	hypothetical prote
29	42.5	37.6	3944	2 T19997	hypothetical prote

30	42	37.2	128	2 S76955	hypothetical prote
31	42	37.2	132	2 B82824	protein-export mem
32	42	37.2	184	2 T35841	probable membrane
33	42	37.2	216	2 H82508	transcription regu
34	42	37.2	221	2 A75366	probable transcrip
35	42	37.2	232	2 G75608	hypothetical prote
36	42	37.2	311	2 S66600	cytochrome-c oxida
37	42	37.2	364	2 B64766	YaiW protein - Esc
38	42	37.2	367	2 G83309	conserved hypothet
39	42	37.2	459	2 T43558	sulfide dehydrogen
40	42	37.2	512	2 T40615	probable ubiquitin
41	42	37.2	546	2 A32260	cholesterol oxidase
42	42	37.2	547	2 T48551	fructosidase-like
43	42	37.2	581	2 A42743	pol polyprotein -
44	42	37.2	677	2 T07944	probable long-chain
45	42	37.2	843	1 GNVWK	pol polyprotein -

ALIGNMENTS

RESULT 1
T47959
hypothetical protein F15G16.60 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #extl_change 20-Apr-2000
C:Accession: T47959
R:De Haan, M.; Maree, A.C.; Grievell, L.A.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.: Q
submitted to the Protein Sequence Database, January 2000
A:Reference number: Z24480
A:Accession: T47959
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-790 <DB>
A:Cross-references: EMBL:AL132959
A:Experimental source: cultivar Columbia; BAC clone F15G16
C:Genetics:
A:Map position: 3
A:Introns: 39/1: 678/2; 698/3; 773/2
A:Note: F15G16.60

Query Match 47.8%; Score 54; DB 2; Length 790;
Best local Similarity 52.6%; Pred. No. 1.9;
Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 TPDINPAMYASRGIRPVGRF 20
DB 366 PPHNPRYSGRGIRPVGRF 384

RESULT 2
F83376
conserved hypothetical protein PA2151 [imported] - Pseudomonas aeruginosa (strain PAO
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #extl_change 31-Dec-2000
R:Accession: F83376
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Laidig, K.; L
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pa
A:Reference number: AB2950; NCID:20437337
A:Accession: F83376
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-664 <STO>
A:Cross-references: GB:AE004642; GB:AE004091; NID:99948163; PTDN:AA05539.1; GSPDB:GN
C:Experimental source: strain PAO1
C:Genetics: PA2151
A:Gene: PA2151

Query Match 44.7%; Score 50.5; DB 2; Length 664;
 Best Local Similarity 62.5%; Pred. No. 5.5;
 Matches 10; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

QY 1 TPDINPAMYASRGIRP 16
 ||||| : : ||
 DB 478 TPDINP-WFQSGSRP 492

RESULT 3
 F71015
 hypothetical protein PH1420 - Pyrococcus horikoshii

C:Species: Pyrococcus horikoshii
 C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 20-Jun-2000
 C:Accession: F71015
 R:Kawarabayashi, Y.; Sawada, N.; Horikawa, H.; Hakiwa, Y.; Hino, Y.; Yamamoto, S.; Sekita, N.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kishida, N.; Oguchi, N. Res. 5, 55-76, 1998
 A>Title: Complete sequence and gene organization of the genome of a hyper-thermophilic
 A:Reference number: A71000; MUID:98344137
 A:Accession: F71015
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-284 <KAW>
 A:Cross-references: GB:AF000006; NID:g3236133; PLIN:BAA30526.1; PID:g3257843
 A:Experimental source: strain ON3
 A:Note: this accession replaces an interim accession for a sequence replaced by Genbank
 C:Genetics:
 A:Gene: PH1420
 C:Superfamily: Pyrococcus horikoshii hypothetical protein PH1420

Query Match 41.6%; Score 47; DB 2; Length 284;
 Best Local Similarity 50.0%; Pred. No. 8;
 Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 PDINPAMYASRGIRPVG 19
 | | | | | : : ||
 DB 217 PYIEPTFVALKGLDLGR 234

RESULT 4

H82852
 hydroxybenzoate octaprenyltransferase XF0068 [imported] - Xylella fastidiosa (strain 9a5
 C:Species: Xylella fastidiosa
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 17-Nov-2000
 C:Accession: H82852
 R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
 Nature 406, 151-157, 2000
 A>Title: The genome sequence of the plant pathogen Xylella fastidiosa.
 A:Reference number: A82515; MUID:20365717
 A:Note: for a complete list of authors see reference number A59328 below
 A:Accession: H82852
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-333 <SIN>
 A:Cross-references: GB:AE003860; GB:AE003849; NID:g9104830; PLIN:AF82881.1; GSPDB:GN001
 A:Experimental source: strain 9a5c
 R:Simpton, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, N.; Alvarenga, R.; A
 Briones, M.R.S.; Bueno, N.R.P.; Camargo, L.E.A.; Camargo, D.N.; Carier, H
 A:Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
 submitted to Genbank, June 2000
 A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Froim
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kurame, E.E.; Laig
 Chado, M.A.; Madeira, A.N.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
 A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Maracca, E.C.; Miyaki, C.Y.;
 F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Savasak
 A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.N.; Silva Jr., W.A.; da Silveir
 M.; Tsubako, N.H.; Valada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
 A:Reference number: A59328
 A:Content type: annotation
 C:Genetics:

A:Gene: XF0068
 C:Superfamily: 4-hydroxybenzoate octaprenyltransferase

Query Match 41.6%; Score 47; DB 2; Length 333;
 Best Local Similarity 53.3%; Pred. No. 9.4;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 4 INPAMYASRGIRPVG 18
 : | | : | | | | |
 DB 54 LDPYKILARDRPG 68

RESULT 5

T49717
 related to BCS1 protein precursor [imported] - Neurospora crassa
 N:Alternate names: protein B23L21.300
 C:Species: Neurospora crassa
 C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
 C:Accession: T49717
 R:Schulte, U.; Aign, V.; Hobeisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatu
 submitted to the Protein Sequence Database, May 2000
 A:Reference number: Z25022
 A:Accession: T49717
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-779 <SCH>
 A:Cross-references: EMBL:ALJ56172; GSPDB:GN00116; NCSP:B23L21.300
 A:Experimental source: BAC clone B23L21; strain OR74A
 C:Genetics:
 A:Gene: NCSP:B23L21.300
 A:Map position: 6

Query Match 41.2%; Score 46.5; DB 2; Length 779;
 Best Local Similarity 58.8%; Pred. No. 27;
 Matches 10; Conservative 2; Mismatches 2; Indels 3; Gaps 1;

QY 1 TPDINPA---NYSRGI 14
 | : | | | : | | : | |
 DB 286 TDYINPATRRRYANRGI 302

RESULT 6

T16859
 hypothetical protein T13C2.5 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Nov-2000
 C:Accession: T16859
 R:Du, Z.
 submitted to the EMBL Data Library, November 1995
 A:Description: The sequence of C. elegans cosmid T13C2.
 A:Reference number: Z16591
 A:Accession: T16859
 A:Status: preliminary; translated from GB/ENBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1296 <DUZ>
 A:Cross-references: EMBL:U40030; NID:g1055164; PID:g1055165; PLIN:AA81133.1; CESP:T1
 C:Genetics:
 A:Gene: CESP:T13C2.5
 A:Introns: 62/3; 93/3; 124/3; 142/1; 212/1; 293/1; 336/2; 406/3; 437/1; 508/1; 584/1;

Query Match 40.7%; Score 46; DB 2; Length 1296;
 Best Local Similarity 50.0%; Pred. No. 54;
 Matches 10; Conservative 2; Mismatches 6; Indels 2; Gaps 1;

QY 1 TPDINPAMYASRGIRPVG 20
 | | | | | : : : : |
 DB 975 TTDINSDFRSR--KDIINF 992

RESULT 7

A>Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
 A:Reference number: A70500; MUID:98295987
 A:Accession: D70817
 A>Status: Preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-276 <COL>
 A:Cross-references: GB:AL022003; GB:AL123456; NID:g3261547; PIDN:CMAL7613.1; PID:g326154
 A:Experimental source: strain H37Rv
 C:Genetics:
 A:Gene: RV1716

Query Match 38.9%; Score 44; DB 2; Length 276;
 Best Local Similarity 20.0%; Pred. No. 23;
 Matches 10; Conservative 5; Mismatches 3; Indels 32; Gaps 1;

OY 1 TPPIINP-----WYSRGRIPYG 18
 |||||
 DB 111 TPDIRGDIYVNTGMHRYADSERYAYSPGFDKKGEMFAKGVKAVG 160

RESULT 12
 A75117
 A:hypothetical protein PAB0502 - *Pyrococcus abyssi* (strain Orsay)
 C:Species: *Pyrococcus abyssi*
 C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
 C:Accession: A75117
 R:anonymous, Genoscope
 submitted to the EMBL Data Library, July 1999
 A:Description: *Pyrococcus abyssi* genome sequence: insights into archaeal chromosome stry
 A:Reference number: A75001
 A:Accession: A75117
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-284 <KAN>
 A:Cross-references: GB:AJ248285; GB:AL096836; NID:g5458067; PIDN:CAB49650.1; PID:g545816
 A:Experimental source: strain Orsay
 C:Genetics:
 A:Gene: PAB0502
 C:Superfamily: *Pyrococcus horikoshii* hypothetical protein PH1420

Query Match 38.9%; Score 44; DB 2; Length 284;
 Best Local Similarity 44.4%; Pred. No. 23;
 Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
 OY 2 PDINPAMVASKRGIRPYGR 19
 |||||
 DB 217 PYIEPTVALRGLELLGK 234

RESULT 13
 B70631
 A:hypothetical protein Rv0428c - *Mycobacterium tuberculosis* (strain H37Rv)
 C:Species: *Mycobacterium tuberculosis*
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
 C:Accession: B70631
 R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skellon, S.; Squares, S.
 Nature 393, 537-544, 1998
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A>Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
 A:Reference number: A70500; MUID:98295987
 A:Accession: B70631
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-302 <COL>
 A:Cross-references: GB:284724; GB:AL123456; NID:g3261708; PIDN:CAB06568.1; PID:g1817694
 A:Experimental source: strain H37Rv
 C:Genetics:
 A:Gene: Rv0428c
 C:Superfamily: *Mycobacterium tuberculosis* hypothetical protein Rv0428c

Query Match 38.9%; Score 44; DB 2; Length 302;
 Best Local Similarity 70.0%; Pred. No. 25;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 OY 7 AMVASRGIRP 16
 |||||
 DB 126 AMVASRDLP 135

RESULT 14
 T32376
 A:hypothetical protein K10F12.4 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T32376
 R:Wohlmann, P.; Beck, C.
 submitted to the EMBL Data Library, September 1997
 A:Description: The sequence of *C. elegans* cosmid K10F12.
 A:Reference number: 221157
 A:Accession: T32376
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-309 <MOCH>
 A:Cross-references: EMBL:AF025462; PIDN:AA871002.1; GSPDB:GM00021; CESP:K10F12.4
 A:Experimental source: strain Bristol N2; clone K10F12
 C:Genetics:
 A:Gene: CESP:K10F12.4
 A:Map position: 3
 A:introns: 31/3; 123/2; 196/3; 239/1

Query Match 38.9%; Score 44; DB 2; Length 309;
 Best Local Similarity 44.4%; Pred. No. 25;
 Matches 8; Conservative 3; Mismatches 5; Indels 2; Gaps 1;
 OY 2 PDINPAMVASKRGIRPYGR 19
 |||||
 DB 131 PDSPNMTLPKS--PIGR 146

RESULT 15
 F81694
 A:pyruvate dehydrogenase, E2 component, dihydrolipoamide acetyltransferase, probable TC
 C:Species: *Chlamydia muridarum*, *Chlamydia trachomatis* MOPn
 C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-May-2000
 C:Accession: F81694
 R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke
 C.; Dodson, R.; Gwinn, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzbe
 Nucleic Acids Res. 28, 1397-1406, 2000
 A:Title: Genome sequences of *Chlamydia trachomatis* MOPn and *Chlamydia pneumoniae* AR39
 A:Reference number: A81500; MUID:20150253
 A:Accession: F81694
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-428 <RET>
 A:Cross-references: GB:AE002320; GB:AE002160; NID:g7190551; PIDN:AAF39360.1; PID:g719
 A:Experimental source: strain N199 (MOPn)
 C:Genetics:
 A:Gene: TC0518
 C:Superfamily: dihydrolipoamide acetyltransferase; lipoyl/biotin-binding homology

Query Match 38.9%; Score 44; DB 2; Length 428;
 Best Local Similarity 37.5%; Pred. No. 36;
 Matches 6; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
 OY 2 PDINPAMVASKRGIRPV 17
 |||||
 DB 196 PDVNPAGSYVESISPI 211

Tue Apr 17 15:46:21 2001

us-09-446-543a-64.rpr

Page 5

Search completed: April 17, 2001, 15:45:58
Job time: 606 sec

```

RESULT 2
PRRP_BOVIN STANDARD: PRT: 98 AA.
AC P81264:
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
DE PROLACTIN-RELEASING PEPTIDE PRECURSOR (PRRP) (PROLACTIN-RELEASING
DE HORMONE) [CONTAINS: PROLACTIN-RELEASING PEPTIDE PRRP1; PROLACTIN-
DE RELEASING PEPTIDE PRRP20].
GN PRH.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM M.A., AND SEQUENCE OF 23-52.
RC TISSUE=Brain;
RX MEDLINE=98268781; PubMed=9607765;
RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,
RA Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,
RA Kurokawa T., Nishimura O., Onda H., Fujino M.;
RT "A prolactin-releasing peptide in the brain.";
RL Nature 393:272-276(1998).
CC -1- FUNCTION: STIMULATES PROLACTIN (PRL) RELEASE AND REGULATES THE
CC EXPRESSION OF PROLACTIN THROUGH ITS RECEPTOR HGR3. MAY STIMULATE
CC LACTOTROPHS DIRECTLY TO SECRETE PRL.
CC -1- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC -----
DR EMBL: AB015417; BAA29025.1; -
KW Hormone; Amidation; Signal; Cleavage on pair of basic residues.
FT SIGNAL 1 22
FT PEPTIDE 23 53 PROLACTIN-RELEASING PEPTIDE PRRP31.
FT PEPTIDE 33 53 PROLACTIN-RELEASING PEPTIDE PRRP20.
FT MOD_RES 53 53 AMIDATION (G-54 PROVIDE AMIDE GROUP).
SQ SEQUENCE 98 AA; 10544 MW; 08AC35A13B0FA908 CRC64;

Query Match 96.5%; Score 109; DB 1; Length 98;
Best Local Similarity 95.0%; Pred. No. 2.2e-10;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPIDINPAMYASRGIRPVGRF 20
DB 34 TPIDINPAMYAGRGIRPVGRF 53

RESULT 3
PRRP_RAT STANDARD: PRT: 83 AA.
AC P81278;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
DE PROLACTIN-RELEASING PEPTIDE PRECURSOR (PRRP) (PROLACTIN-RELEASING
DE HORMONE) [CONTAINS: PROLACTIN-RELEASING PEPTIDE PRRP1; PROLACTIN-
DE RELEASING PEPTIDE PRRP20].
GN PRH.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC TISSUE=Brain;
RX MEDLINE=98268781; PubMed=9607765;
RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,
RA Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,
RA Kurokawa T., Nishimura O., Onda H., Fujino M.;
RT "A prolactin-releasing peptide in the brain.";
RL Nature 393:272-276(1998).
CC -1- FUNCTION: STIMULATES PROLACTIN (PRL) RELEASE AND REGULATES THE
CC EXPRESSION OF PROLACTIN THROUGH ITS RECEPTOR HGR3. MAY STIMULATE
CC LACTOTROPHS DIRECTLY TO SECRETE PRL.
CC -1- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC -----
DR EMBL: AB015418; BAA29026.1; -
KW Hormone; Amidation; Signal; Cleavage on pair of basic residues.
FT SIGNAL 1 21
FT PEPTIDE 22 52 PROLACTIN-RELEASING PEPTIDE PRRP31.
FT PEPTIDE 33 52 PROLACTIN-RELEASING PEPTIDE PRRP20.
FT MOD_RES 52 52 AMIDATION (G-53 PROVIDE AMIDE GROUP).
SQ SEQUENCE 83 AA; 9215 MW; D0C75A264EEB4F29 CRC64;

Query Match 92.9%; Score 105; DB 1; Length 83;
Best Local Similarity 90.0%; Pred. No. 7.8e-10;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TPIDINPAMYASRGIRPVGRF 20
DB 33 TPIDINPAMYTGIRPVGRF 52

RESULT 4
EXTL_HUMAN STANDARD: PRT: 676 AA.
AC Q92935;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
DE EXOSTOSIN-LIKE 1 (EXOSTOSIN-L) (MULTIPLE EXOSTOSIS-LIKE PROTEIN).
GN EXTL1 OR EXTL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97189339; PubMed=9037597;
RA Wise C.A., Clines G.A., Massa H., Trask B.J., Lovett M.;
RT "Identification and localization of the gene for EXTL1, a third member
RT of the multiple exostoses gene family.";
RL Genome Res. 7:10-16(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Xu L., Deng H.X., Xia J.H., Pan Q., Liu C.Y.;
RT "Mutations of the EXT genes in hereditary multiple exostoses in
RT Chinese.";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Huys W., Spleker N., Van Roy N., De Paeppe A., De Bouille K.,
RA Williams P.J., Van Hul W., Versteeg R., Speleman F.;
RT "Refined physical mapping and genomic structure of the EXTL1 gene.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: APPEARS TO BE A TUMOR SUPPRESSOR (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ENDOPLASMIC
CC RETICULUM (BY SIMILARITY).

```

CC -1- SIMILARITY: BELONGS TO THE EXOSTOSIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U67191; AAC51141.1; -
 CC EMBL: AF083633; AAD02840.1; -
 CC EMBL: AF083623; AAD02840.1; JOINED.
 CC EMBL: AF083624; AAD02840.1; JOINED.
 CC EMBL: AF083625; AAD02840.1; JOINED.
 CC EMBL: AF083626; AAD02840.1; JOINED.
 CC EMBL: AF083627; AAD02840.1; JOINED.
 CC EMBL: AF083628; AAD02840.1; JOINED.
 CC EMBL: AF083629; AAD02840.1; JOINED.
 CC EMBL: AF083630; AAD02840.1; JOINED.
 CC EMBL: AF083631; AAD02840.1; JOINED.
 CC EMBL: AF083632; AAD02840.1; JOINED.
 CC EMBL: AF153980; AAF73172.1; -
 CC EMBL: AF15391; AAF73172.1; JOINED.
 CC MIM: 601738; -
 CC Anti-oncogene; Multigene family: Transmembrane; Signal-anchor.
 CC TRANSMEM 10 30 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
 CC FT CAROHW 269 269 (POTENTIAL).
 CC FT SEQUENCE 676 AA; 74673 MW; B5E06A8762E5633 CRC64;
 CC SQ

Query Match 41.6%; Score 47; DB 1; Length 676;
 Best Local Similarity 45.0%; Pred. No. 7.8;
 Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
 QY 1 TPDIINPAWYASRGIRPVGR 20
 Db 400 SPQDPFYTLQGSREGRF 419

RESULT 5
 CYCR_CHRY1 STANDARD; PRT; 383 AA.
 AC 082947;
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE PHOTOSYNTHETIC REACTION CENTER CYTOCHROME C SUBUNIT PRECURSOR.
 GN PUPC.
 OS Chromatium vinosum.
 OC Bacteria; Proteobacteria; gamma subdivision; Chromatiaceae;
 OC Allochroa.
 OX NCBI_TaxID=1049;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=D;
 RA Corson G.E., Nagashima K.V., Matsuura K., Sakuragi Y., Ruwanchi W.,
 RA Qin N., Allen R., Knaf D.B.;
 RA "Primary structure of genes encoding light-harvesting and reaction
 RT center proteins from Chromatium vinosum."
 RT Submitted (MAR-1998) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: THE REACTION CENTER OF PURPLE BACTERIA CONTAIN A
 CC TIGHTLY BOUND CYTOCHROME MOLECULE WHICH RE-REDUCES THE PHOTO
 CC OXIDIZED PRIMARY ELECTRON DONOR.
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A LIPID
 CC ANCHOR (BY SIMILARITY).
 CC -1- PTM: BINDS FOUR HEME GROUPS PER MOLECULE.
 CC -1- SIMILARITY: HIGH, WITH OTHER PHOTOSYNTHETIC REACTION CENTER
 CC CYTOCHROME C SUBUNITS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AB011811; BAA32742.1; -
 CC InterPro: IPR000345; -
 CC DR PROSITE: PS00190; CYTOCHROME C; 4.
 CC Electron transport; Photosynthesis; Reaction center; Heme;
 CC Membrane; Lipoprotein; Duplication; Signal.
 CC FT SIGNAL 1 22
 CC FT CHAIN 23 383
 CC FT LIPID 23 23
 CC FT BINDING 107 107
 CC FT BINDING 110 110
 CC FT METAL 111 111
 CC FT BINDING 152 152
 CC FT BINDING 155 155
 CC FT METAL 156 156
 CC FT BINDING 247 247
 CC FT BINDING 250 250
 CC FT METAL 251 251
 CC FT BINDING 307 307
 CC FT BINDING 310 310
 CC FT METAL 311 311
 CC SQ SEQUENCE 383 AA; 41522 MW; 96BCD91EF1B9A87E CRC64;
 CC FT

Query Match 40.7%; Score 46; DB 1; Length 383;
 Best Local Similarity 42.1%; Pred. No. 6.3;
 Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
 QY 1 TPDIINPAWYASRGIRPVGR 19
 Db 263 TPQRTAMATAIHVNDINQ 281

RESULT 6
 YAAJ_ECOLI STANDARD; PRT; 476 AA.
 AC P30143;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE HYPOTHETICAL 51.7 KDA PROTEIN IN THRC-TALB INTERGENIC REGION (ORF8).
 GN YAAJ.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RC MEDLINE=92334977; PubMed=1630901;
 RA Yura T., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N.,
 RA Isono K., Mizubuchi K., Nakata A.;
 RT "Systematic sequencing of the Escherichia coli genome: analysis of
 RT the 0-2.4 min region."
 RL Nucleic Acids Res. 20:3305-3308(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Colado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;

RI "The complete genome sequence of *Escherichia coli* K-12.";
 RL Science 277:1453-1474(1997).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
 CC (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE SODIUM:ALANINE SYMPORTER FAMILY
 CC (SAF). STRONG, TO H. INFLUENZAE N10183.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: D10483; -, NOT_ANNOTATED_CDS.
 DR EMBL: AE000111; AAC73118.1; -.
 DR Ecocyc: Ecol1555; yaad.
 DR InterPro: IPR001463; -.
 DR Pfam: PF01235; Na_Ala_symp. 1.
 DR PRINTS: PR00175; NALALASPORT.
 DR PROSITE: PS00873; NA_ALANINE_SYMP. 1.
 DR Nucleotide: P000175; NALALASPORT.
 KM Nucleotide: P000175; NALALASPORT.
 FT TRANSMEM 4 24 POTENTIAL.
 FT TRANSMEM 81 101 POTENTIAL.
 FT TRANSMEM 141 161 POTENTIAL.
 FT TRANSMEM 174 194 POTENTIAL.
 FT TRANSMEM 207 227 POTENTIAL.
 FT TRANSMEM 233 253 POTENTIAL.
 FT TRANSMEM 300 320 POTENTIAL.
 FT TRANSMEM 351 371 POTENTIAL.
 FT TRANSMEM 391 411 POTENTIAL.
 FT TRANSMEM 414 434 POTENTIAL.
 SQ SEQUENCE 476 AA; 51662 MW; 256B2E12E126B53 CRC64;

Query Match 38.9%; Score 44; DB 1; Length 476;
 Best Local Similarity 44.4%; Pred. NO. 16;
 Matches 8; Conservative 3; Mismatches 1; Indels 6; Gaps 1;

QY 3 DIN-----PANYASRG1 14
 1:1 1111:11:
 DB 120 DVMGQFRGSPAWYMARGL 137

RESULT 7
 ARP_YEAST STANDARD; PRT; 719 AA.
 ID ARP_YEAST 012228;
 AC P32770; 012228;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE ARP PROTEIN.
 GN ARP OR ARP OR YDL167C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 CX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AN22;
 RX MEDLINE=93247548; PubMed=8483449;
 RA Wehner E.P., Rao E., Brendel M.;
 RT "Molecular structure and genetic regulation of SFA, a gene
 RT responsible for resistance to formaldehyde in *Saccharomyces*
 RT *cerevisiae*, and characterization of its protein product.";
 RL Mol. Gen. Genet. 237:351-358(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RA Pohl T.M.;
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: CONTAINS 2 RANBP-TYPE ZINC FINGERS.
 CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: X68020; CAA48159.1; -.
 DR EMBL: Z67750; CAA91579.1; -.
 DR EMBL: Z74215; CAA98741.1; -.
 DR PIR: S31139; S31139.
 DR NSSP: P04170; 6RXM.
 DR SGD: S0002326; NRPL.
 DR InterPro: IPR000504; -.
 DR InterPro: IPR001876; -.
 DR Pfam: PF00076; rrm; 1.
 DR Pfam: PF00641; zf-RanBP; 2.
 DR PROSITE: PS50102; RRM; 1.
 DR PROSITE: PS00030; RRM_RNP_1; FALSE_NEG.
 DR Nucleotide: P000175; NALALASPORT.
 FT DOMAIN 226 322 RNA-BINDING (RRM).
 FT DOMAIN 490 564 ASN-RICN.
 FT COMPLET 493 493 I -> N (IM REF. 1).
 SQ SEQUENCE 719 AA; 79299 MW; ADA9BC09FD582659 CRC64;

Query Match 38.9%; Score 44; DB 1; Length 719;
 Best Local Similarity 50.0%; Pred. NO. 24;
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 7 AMYASRG1RPG 18
 1:1 11111:
 DB 244 SWFTQYGVPRVG 255

RESULT 8
 VG12_BPBO3 STANDARD; PRT; 860 AA.
 ID VG12_BPBO3 037893;
 AC Q37893;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE PRE-NECK APPENDAGE PROTEIN (LATE PROTEIN GP12).
 GN Bacteriophage B103.
 OS Viruses; dsDNA viruses, no RNA stage; Tailed phages; Podoviridae.
 CC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Podoviridae.
 CX NCBI_TaxID=10778;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96019084; PubMed=9358052;
 RA Pachenkova T., Benes V., Paces V., Vlcek C., Paces V.;
 RT Bacteriophage B103: complete DNA sequence of its genome and
 RT relationship to other Bacillus phages.";
 RL Gene 199:157-163(1997).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: X99260; CAA67660.1; -.
 DR EMBL: X99260; CAA67660.1; -.
 KM Late protein.
 SQ SEQUENCE 860 AA; 91741 MW; ACSB38F4BC65ECB0 CRC64;

Query Match 38.5%; Score 43.5; DB 1; Length 860;

RT Nucleotide sequence of the gene for cholesterol oxidase from a
 RT Streptomyces sp.":
 RL J. Bacteriol. 171:596-601(1989).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS), AND MUTAGENESIS.
 RX MEDLINE=99211873; PubMed=10194345;
 RA Yue Q.K., Kass I.J., Sampson N.S., Vrielink A.;
 RT Crystal structure determination of cholesterol oxidase from
 RT Streptomyces and structural characterization of key active site
 RT mutants":
 RL Biochemistry 38:4277-4286(1999).
 CC -1- CATALYTIC ACTIVITY: CHOLESTEROL + O(2) -> CHOLEST-4-EN-3-ONE +
 CC H(2)O(2).
 CC -1- COFACTOR: FAD.
 CC -1- PATHWAY: CHOLESTEROL METABOLISM.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- SIMILARITY: BELONGS TO THE GMC OXIDOREDUCTASES FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 CC EMBL: M31939; AAA26719.1; -
 CC PIR: A32260; A32260.
 DR PDB: 1B4V; 06-JAN-99.
 DR PDB: 1B8S; 09-FEB-99.
 DR PDB: 1CB0; 10-MAR-99.
 DR PDB: 1CC2; 11-MAR-99.
 DR InterPro: IPR000172; -
 DR InterPro: IPR001167; -
 DR Pfam: PF01319; CHOD; 1.
 DR PROSITE: PS00623; GMC_OXRED-1; 1.
 DR PROSITE: PS00624; GMC_OXRED-2; FALSE_NEG.
 KW Oxidoreductase; Signal; Flavoprotein; FAD; 3D-structure.
 FT SIGNAL 1 42
 FT CHAIN 43 546
 FT NP_BIND 54 70 CHOLESTEROL OXIDASE.
 FT ACT_SITE 398 398 FAD (ADP PART) (POTENTIAL).
 FT ACT_SITE 484 484 PROTON ACCEPTOR.
 FT MUTAGEN 398 398 E->Q: REDUCED ACTIVITY.
 FT MUTAGEN 484 484 H->N,Q: REDUCED ACTIVITY.
 SO SEQUENCE 546 AA; 58993 MW; EF22A1FE5BA68D21 CRC64;

Query Match 37.2%; Score 42; DB 1; Length 546;
 Best Local Similarity 36.8%; Pred. No. 38;
 Matches 7; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 2 PDINPAMYASRGIRPVGR 20
 || : : : : :
 DB 98 PDKRSSWFKNRTPLGSP 116

RESULT 14
 POL_MLVAK STANDARD; PRT; 581 AA.
 AC P31795;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE POL POLYPROTEIN [CONTAINS: PROTEASE (EC 3.4.23.-); REVERSE
 DE TRANSCRIPTASE (EC 2.7.7.49); RIBONUCLEASE H (EC 3.1.26.4)] (FRAGMENT).
 GN POL.
 OS Radiation murine leukemia virus (strain Kaplan).
 OC Viruses; Retroviral viruses; Retroviridae; Mammalian type C retroviruses.
 OX NCBI_TaxID=31689;
 RN [1]
 DR SEQUENCE FROM N.A.

RX MEDLINE=92333703; PubMed=1629969;
 RA Pollquin L., Bergeron D., Fortier J.L., Paquette Y., Bergeron R.,
 RA Rassat E.;
 RT Determinants of thymotropism in Kaplan radiation leukemia virus and
 RT nucleotide sequence of its envelope region.";
 RL J. Virol. 66:5141-5146(1992).
 CC -1- PPM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC -1- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A GAG-POL
 CC POLYPROTEIN.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A2; ALSO KNOWN AS THE
 CC RETROPEPSIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 CC EMBL: M93052; AAA6525.1; -
 CC PIR: A42743; A42743.
 DR MEROPS: A02.008; -
 DR InterPro: IPR001037; -
 DR InterPro: IPR001584; -
 DR InterPro: IPR001969; -
 DR InterPro: IPR002156; -
 DR Pfam: PF00552; Integrase; 1.
 DR Pfam: PF00075; Integrase; 1.
 DR Pfam: PF00665; Ite; 1.
 DR PROSITE: PS00141; ASP_PROTEASE; PARTIAL.
 KW Hydrolyase; Transferase; RNA-directed DNA polymerase;
 KW Aspartyl protease; Endonuclease; Polypeptin.
 FT NON_TER 1 1
 SO SEQUENCE 581 AA; 65157 MW; 8D7A38694C8E036E CRC64;

Query Match 37.2%; Score 42; DB 1; Length 581;
 Best Local Similarity 41.2%; Pred. No. 41;
 Matches 7; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 TPDINPAMYASRGIRPV 17
 || | : | : | :
 DB 556 TPPIRPSGWRVGRSQNP 572

RESULT 15
 POL_MLVAK STANDARD; PRT; 843 AA.
 AC P03357;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE POL POLYPROTEIN [CONTAINS: REVERSE TRANSCRIPTASE (EC 2.7.7.49);
 DE ENDONUCLEASE] (FRAGMENT).
 GN POL.
 OS AKR murine leukemia virus.
 OC Viruses; Retroviral viruses; Retroviridae; Mammalian type C retroviruses.
 OX NCBI_TaxID=11789;
 RN [1]
 DR SEQUENCE FROM N.A.
 RX MEDLINE=83090450; PubMed=6294621;
 RA Hart W., Corbin V., Gilbert W.;
 RT "Nucleotide sequence of the 3' half of AKV.";
 RL Nucleic Acids Res. 10:6931-6944(1982).
 CC -1- PPM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 DR PIR: A03958; GAVWK.
 DR HSSP: P03355; IMUL.
 DR InterPro: IPR001037; -
 DR InterPro: IPR001584; -
 DR InterPro: IPR002156; -
 DR Pfam: PF00552; Integrase; 1.
 DR Pfam: PF00075; Integrase; 1.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw Model

Run on: April 17, 2001, 15:39:48 ; Search time 61.54 seconds
(without alignments)
6.243 Million cell updates/sec

Title: US-09-446-543a-64
Perfect score: 113
Sequence: 1 TPDINPMATASRGIRPVGRF 20

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
- 3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
- 4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
- 5: /cgn2_6/ptodata/2/1aa/PCTUS.COMB.pep:*
- 6: /cgn2_6/ptodata/2/1aa/Backfile1.pep:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	113	100.0	20	US-09-105-678A-46	Sequence 46, Appl
2	113	100.0	21	US-09-105-678A-47	Sequence 47, Appl
3	113	100.0	22	US-09-105-678A-48	Sequence 48, Appl
4	113	100.0	31	US-09-105-678A-9	Sequence 9, Appl
5	113	100.0	31	US-09-105-678A-43	Sequence 43, Appl
6	113	100.0	32	US-09-105-678A-44	Sequence 44, Appl
7	113	100.0	33	US-09-105-678A-45	Sequence 45, Appl
8	109	96.5	20	US-09-105-678A-34	Sequence 34, Appl
9	109	96.5	21	US-09-105-678A-35	Sequence 35, Appl
10	109	96.5	22	US-09-105-678A-36	Sequence 36, Appl
11	109	96.5	31	US-09-105-678A-7	Sequence 7, Appl
12	109	96.5	31	US-09-105-678A-31	Sequence 31, Appl
13	109	96.5	32	US-09-105-678A-32	Sequence 32, Appl
14	109	96.5	33	US-09-105-678A-33	Sequence 33, Appl
15	109	96.5	20	US-09-105-678A-40	Sequence 40, Appl
16	109	96.5	21	US-09-105-678A-28	Sequence 28, Appl
17	109	96.5	21	US-09-105-678A-41	Sequence 41, Appl
18	109	96.5	22	US-09-105-678A-42	Sequence 42, Appl
19	109	96.5	31	US-09-105-678A-8	Sequence 8, Appl
20	109	96.5	31	US-09-105-678A-37	Sequence 37, Appl
21	109	96.5	31	US-09-105-678A-38	Sequence 38, Appl
22	109	96.5	32	US-09-105-678A-39	Sequence 39, Appl
23	109	96.5	33	US-09-105-678A-30	Sequence 30, Appl
24	109	96.5	29	US-09-105-678A-29	Sequence 29, Appl
25	109	96.5	349	US-08-118-270-71	Sequence 71, Appl
26	109	96.5	349	PCT-US93-08528-71	Sequence 71, Appl
27	109	96.5	349	PCT-US93-08528-71	Sequence 71, Appl

28	42	37.2	547	1	US-08-083-948-8	Sequence 8, Appl
29	42	37.2	547	1	US-08-393-785-8	Sequence 8, Appl
30	42	37.2	547	1	US-08-475-694-8	Sequence 8, Appl
31	42	37.2	547	1	US-08-712-057-8	Sequence 8, Appl
32	40	35.4	433	1	US-08-417-492-2	Sequence 2, Appl
33	40	35.4	555	2	US-08-982-232-7	Sequence 7, Appl
34	40	35.4	555	2	US-08-982-232-14	Sequence 14, Appl
35	39.5	35.0	831	1	US-08-073-384C-5	Sequence 5, Appl
36	39.5	35.0	831	1	US-08-254-359A-5	Sequence 5, Appl
37	39.5	35.0	831	1	US-08-483-043-5	Sequence 5, Appl
38	39.5	35.0	831	1	US-08-481-238-5	Sequence 5, Appl
39	39.5	35.0	831	2	US-08-471-066B-5	Sequence 5, Appl
40	39.5	35.0	831	2	US-08-484-956-5	Sequence 5, Appl
41	39.5	35.0	831	2	US-08-757-653-5	Sequence 5, Appl
42	39.5	35.0	831	2	US-08-599-491-5	Sequence 5, Appl
43	39.5	35.0	831	2	US-08-756-386-5	Sequence 5, Appl
44	39.5	35.0	831	2	US-08-823-516-5	Sequence 5, Appl
45	39.5	35.0	831	3	US-08-682-853A-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-09-105-678A-46
Sequence 46, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESSES:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09105, 678A
FILING DATE: 26-JUN-1998
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-46

Query Match 100.0%; Score 113; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 1e-11; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0;

Qy 1 TPDINPAMYASRGIRPVGRF 20
|||||
Db 1 TPDINPAMYASRGIRPVGRF 20

RESULT 2

US-09-105-678A-47
; Sequence 47, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,678A
; FILING DATE: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-105-678A-47

Query Match 100.0%; Score 113; DB 3; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.1e-11;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPDINPAMYASRGIRPVGRF 20
|||||
Db 1 TPDINPAMYASRGIRPVGRF 20

RESULT 3

US-09-105-678A-48
; Sequence 48, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-48

Query Match 100.0%; Score 113; DB 3; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.1e-11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPDINPAMYASRGIRPVGRF 20
|||||
Db 1 TPDINPAMYASRGIRPVGRF 20

RESULT 4

US-09-105-678A-9
; Sequence 9, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,678A
; FILING DATE: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.

REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-9

Query Match 100.0%; Score 113; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.7e-11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDIINPAWYASRGIRPVGRF 20
|||||
DB 12 TPDIINPAWYASRGIRPVGRF 31

RESULT 5
US-09-105-678A-43
Sequence 43, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-43

Query Match 100.0%; Score 113; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.7e-11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDIINPAWYASRGIRPVGRF 20
|||||
DB 12 TPDIINPAWYASRGIRPVGRF 31

RESULT 6
US-09-105-678A-44
Sequence 44, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 32 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-44

Query Match 100.0%; Score 113; DB 3; Length 32;
Best Local Similarity 100.0%; Pred. No. 1.7e-11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDIINPAWYASRGIRPVGRF 20
|||||
DB 12 TPDIINPAWYASRGIRPVGRF 31

RESULT 7
US-09-105-678A-45
Sequence 45, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:

ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: ConlIn, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-45

Query Match 100.0%; Score 113; DB 3; Length 33;
Best Local Similarity 100.0%; Pred. No. 1,8e-11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAMVASRGTRPVGRF 20
DB 12 TPDINPAMVASRGTRPVGRF 31

RESULT 8
US-09-105-678A-34
Sequence 34, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:

NAME: ConlIn, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-34

Query Match 96.5%; Score 109; DB 3; Length 20;
Best Local Similarity 95.0%; Pred. No. 4,4e-11;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPDINPAMVASRGTRPVGRF 20
DB 1 TPDINPAMVASRGTRPVGRF 20

RESULT 9
US-09-105-678A-35
Sequence 35, Application US/09105678A
Patent No. 6103882

GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: ConlIn, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-35

Query Match 96.5%; Score 109; DB 3; Length 21;
Best Local Similarity 95.0%; Pred. No. 4,6e-11;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPDINPAMYASRGIRPVGRF 20
|||||
Db 1 TPDINPAMYAGRGIRPVGRF 20

RESULT 10
US-09-105-678A-36
; Sequence 36, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,678A
; FILING DATE: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-105-678A-36

Query Match 96.5%; Score 109; DB 3; Length 22;
Best Local Similarity 95.0%; Pred. No. 4.9e-11;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPDINPAMYASRGIRPVGRF 20
|||||
Db 1 TPDINPAMYAGRGIRPVGRF 20

RESULT 11
US-09-105-678A-7
; Sequence 7, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52

CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,678A
; FILING DATE: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-105-678A-7

Query Match 96.5%; Score 109; DB 3; Length 31;
Best Local Similarity 95.0%; Pred. No. 7.2e-11;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPDINPAMYASRGIRPVGRF 20
|||||
Db 12 TPDINPAMYAGRGIRPVGRF 31

RESULT 12
US-09-105-678A-31
; Sequence 31, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,678A
; FILING DATE: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997

ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27, 026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-31

Query Match 96.5%; Score 109; DB 3; Length 31;
Best Local Similarity 95.0%; Pred. No. 7.7e-11;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPDINPAMYASRGIRPVGRF 20
|||||
DB 12 TPDINPAMYAGRGIRPVGRF 31

RESULT 13
US-09-105-678A-32
Sequence 32, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27, 026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 32 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-32

Query Match 96.5%; Score 109; DB 3; Length 32;

Best Local Similarity 95.0%; Pred. No. 7.4e-11;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPDINPAMYASRGIRPVGRF 20
|||||
DB 12 TPDINPAMYAGRGIRPVGRF 31

RESULT 14
US-09-105-678A-33
Sequence 33, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27, 026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-33

Query Match 96.5%; Score 109; DB 3; Length 33;
Best Local Similarity 95.0%; Pred. No. 7.7e-11;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

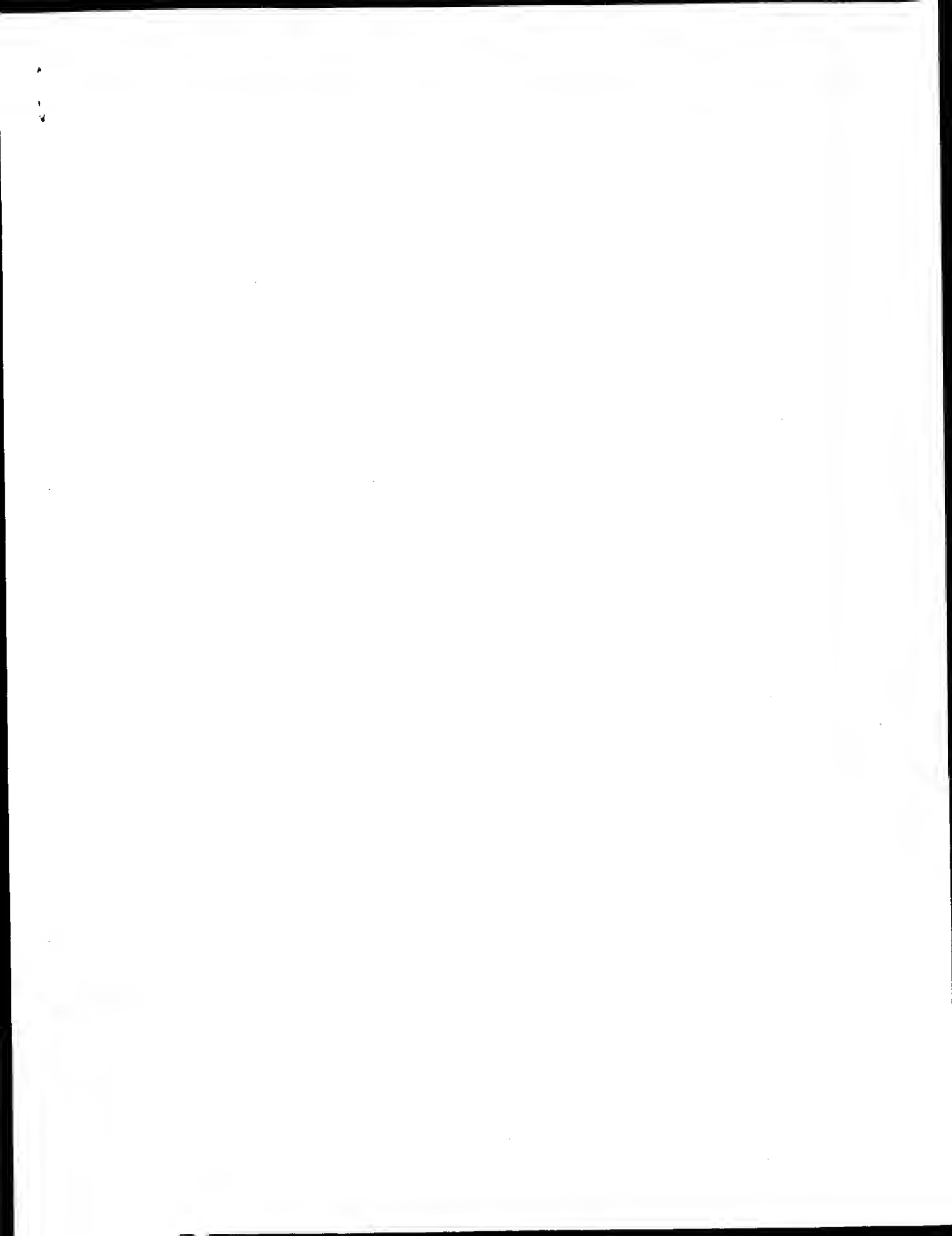
QY 1 TPDINPAMYASRGIRPVGRF 20
|||||
DB 12 TPDINPAMYAGRGIRPVGRF 31

RESULT 15
US-09-105-678A-40
Sequence 40, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND

: NUMBER OF SEQUENCES: 52
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
 : STREET: 130 Water Street
 : CITY: Boston
 : STATE: MA
 : COUNTRY: USA
 : ZIP: 02109
 :
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.30
 :
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/09/105,678A
 : FILING DATE: 26-JUN-1998
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: JP 172118/1997
 : FILING DATE: 27-JUN-1997
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Conlin, David G.
 : REGISTRATION NUMBER: 27,026
 : REFERENCE/DOCKET NUMBER: 48466-342
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 617-523-3400
 : TELEFAX: 617-523-6440
 :
 : INFORMATION FOR SEQ ID NO: 40:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 20 amino acids
 : TYPE: amino acid
 : STRANDEDNESS:
 : TOPOLOGY: linear
 :
 : MOLECULE TYPE: peptide
 :
 : US-09-105-678A-40

Query Match 92.9%; Score 105; DB 3; Length 20;
 Best Local Similarity 90.0%; Pred. No. 1,9e-10;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 TPDINPANYASRGIRPVGRF 20
 ||||| |||||
 DB 1 TPDINPANYTGRGIRPVGRF 20

Search completed: April 17, 2001, 15:39:48
 Job time: 317 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 17, 2001, 15:38:42 ; Search time 116.94 Seconds
(without alignments)
9.777 Million cell updates/sec

Title: US-09-446-543A-64
Perfect score: 113
Sequence: 1 TPDINPAMWASRGIRPVGRF 20

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues
Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_0401:*

- 1: /SIDSL/gcgdata/geneseq/geneseq/AA1980.DAT:*
- 2: /SIDSL/gcgdata/geneseq/geneseq/AA1981.DAT:*
- 3: /SIDSL/gcgdata/geneseq/geneseq/AA1982.DAT:*
- 4: /SIDSL/gcgdata/geneseq/geneseq/AA1983.DAT:*
- 5: /SIDSL/gcgdata/geneseq/geneseq/AA1984.DAT:*
- 6: /SIDSL/gcgdata/geneseq/geneseq/AA1985.DAT:*
- 7: /SIDSL/gcgdata/geneseq/geneseq/AA1986.DAT:*
- 8: /SIDSL/gcgdata/geneseq/geneseq/AA1987.DAT:*
- 9: /SIDSL/gcgdata/geneseq/geneseq/AA1988.DAT:*
- 10: /SIDSL/gcgdata/geneseq/geneseq/AA1989.DAT:*
- 11: /SIDSL/gcgdata/geneseq/geneseq/AA1990.DAT:*
- 12: /SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT:*
- 13: /SIDSL/gcgdata/geneseq/geneseq/AA1992.DAT:*
- 14: /SIDSL/gcgdata/geneseq/geneseq/AA1993.DAT:*
- 15: /SIDSL/gcgdata/geneseq/geneseq/AA1994.DAT:*
- 16: /SIDSL/gcgdata/geneseq/geneseq/AA1995.DAT:*
- 17: /SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT:*
- 18: /SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT:*
- 19: /SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT:*
- 20: /SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT:*
- 21: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT:*
- 22: /SIDSL/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	113	100.0	20	18 W31394	Human type G prote
2	113	100.0	20	20 W97236	Human type ligand
3	113	100.0	20	21 B10365	Human type ligand
4	113	100.0	20	21 Y49294	19P2 ligand peptid
5	113	100.0	21	18 W31395	Human type G prote
6	113	100.0	21	21 B10366	Human type G prote
7	113	100.0	22	18 W31396	Human type G prote
8	113	100.0	22	21 B10367	Human type G prote
9	113	100.0	31	18 W31391	Human type G prote
10	113	100.0	31	20 W97235	Human type ligand
11	113	100.0	31	20 W87615	Human 19P2 ligand

12	113	100.0	31	21 B10362	Human oxytocin sec
13	113	100.0	31	21 Y49291	19P2 ligand peptid
14	113	100.0	32	18 W31392	Human type G prote
15	113	100.0	32	21 B10363	Human oxytocin sec
16	113	100.0	33	18 W31393	Human type G prote
17	113	100.0	33	21 B10364	Human oxytocin sec
18	113	100.0	87	18 W31390	Human type G prote
19	113	100.0	87	20 W97226	Human type ligand
20	113	100.0	87	21 B10361	Human oxytocin sec
21	109	96.5	20	18 W31374	Bovine G protein-c
22	109	96.5	20	20 W97232	Bovine pituitary-d
23	109	96.5	20	20 W95191	Bovine pituitary-d
24	109	96.5	20	21 B10350	Bovine oxytocin se
25	109	96.5	20	21 Y49301	19P2 ligand peptid
26	109	96.5	21	18 W31375	Bovine G protein-c
27	109	96.5	21	20 W95192	Bovine pituitary-d
28	109	96.5	21	21 B10351	Bovine oxytocin se
29	109	96.5	22	18 W31376	Bovine G protein-c
30	109	96.5	22	20 W95193	Bovine pituitary-d
31	109	96.5	22	21 B10352	Bovine oxytocin se
32	109	96.5	31	18 W31371	Bovine G protein-c
33	109	96.5	31	20 W97218	Bovine pituitary-d
34	109	96.5	31	20 W87613	Bovine 19P2 ligand
35	109	96.5	31	20 W95188	Bovine pituitary-d
36	109	96.5	31	21 B10347	Bovine oxytocin se
37	109	96.5	31	21 Y49290	19P2 ligand peptid
38	109	96.5	31	21 Y49298	19P2 ligand peptid
39	109	96.5	32	18 W31372	Bovine G protein-c
40	109	96.5	32	20 W95189	Bovine pituitary-d
41	109	96.5	32	21 B10348	Bovine oxytocin se
42	109	96.5	33	18 W31373	Bovine G protein-c
43	109	96.5	33	20 W95190	Bovine pituitary-d
44	109	96.5	33	21 Y49297	Bovine oxytocin se
45	109	96.5	33	21 Y49297	19P2 ligand peptid

ALIGNMENTS

RESULT	1	
W31394	W31394 standard; Peptide: 20 AA.	
AC	W31394;	
XX		
DT	06-APR-1998 (first entry)	
XX		
DE	Human type G protein-coupled receptor ligand fragment 4.	
XX		
KV	G protein-coupled receptor; ligand binding; pharmaceutical;	
KW	modulator; pituitary; central nervous system; pancreas; prophylactic;	
KV	therapeutic agent.	
XX		
OS	Homo sapiens.	
XX		
PN	W09724436-A2.	
XX		
PD	10-JUL-1997.	
XX		
PF	26-DEC-1996; 96WO-JP03821.	
XX		
PR	18-SEP-1996; 96JP-0246573.	
PR	28-DEC-1995; 95JP-0343371.	
PR	15-MAR-1996; 96JP-0059419.	
PR	12-AUG-1996; 96JP-0211805.	
XX		
PA	(TAKEDA) TAKEDA CHEM IND LTD.	
XX		
PI	Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;	
PI	Kawanata Y, Kitada C;	
XX		
DR	WPI, 1997-363672/33.	
DR	N-PSDB; V02431.	

XX Ligand peptide for G protein-coupled receptor - acts by modulating
 PT function in the central nervous system, pancreas and pituitary gland
 PS
 XX
 PS Claim 2; Page 185; 258pp; English.

CC This sequence represents a peptide fragment from a novel human type
 CC ligand polypeptide corresponding to amino acid residues 34 to 53 of the
 CC sequence represented in W31390 and is used in an assay to monitor ligand
 CC binding to the G protein-coupled receptor protein. Pharmaceutical
 CC compositions containing this ligand may be used as a pituitary function
 CC modulator. A central nervous system modulator or a pancreatic function
 CC modulator. This ligand could have specific applications as a prophylactic
 CC or therapeutic agent for dementia, depression, hyperkinetic syndrome,
 CC disturbance of consciousness, anxiety syndrome, schizophrenia, trauma,
 CC growth hormone secretory disease, hyper- and polyphagia, hyperlipidaemia,
 CC hypercholesterolaemia, hyperglycaemia, hyperprolactinaemia, diabetes,
 CC cancer, pancreatitis, renal disease, Turner's syndrome, neurosis, asthma,
 CC rheumatoid arthritis, spinal injury, transient brain ischaemia, epilepsy,
 CC amyotrophic lateral sclerosis, acute myocardial infarction, infertility,
 CC spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis,
 CC osteoporosis and/or oligosaccharia. Assays can also be developed to screen
 CC compounds which are capable of altering the binding activity of the
 CC ligand affecting activation of the G protein-coupled receptor protein.

XX Sequence 20 AA:

Query Match 100.0%; Score 113; DB 18; Length 20;
 Best Local Similarity 100.0%; Pred. No. 3.5e-12;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TPDIINPAWYASRGIRPGRF 20
 DB 1 tpdiinpaawyasrgirpgrf 20

RESULT 2
 W97236
 ID W97236 standard; peptide: 20 AA.

XX W97236;

XX 06-MAY-1999 (first entry)

DE Human type ligand polypeptide fragment.

XX Rat type ligand; modulation; prolactin secretion;
 KW G protein-coupled receptor; GPCR; hypovarianism; gonocyst cacosgenesis;
 KW menopausal syndrome; euthyroid; hypometabolism; lactation;
 KW pituitary adenomatosis; brain tumour; emmenorrhoea; autoimmune disease;
 KW prolactinoma; infertility; impotence; amenorrhoea; galactorrhoea;
 KW acromegaly; Chlari-Frommel syndrome; Argonz-del Castillo syndrome;
 KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome or dyszoospermia;
 KW contraceptive; placental function; chorioncarcinoma; hydralid mole;
 KW interruption mole; abortion; unfertility fetus; abnormal saccharometabolism;
 KW abnormal lipidmetabolism; oxytocia.

XX Homo sapiens.

XX W09858962-A1.

XX 30-DEC-1998.

XX 22-JUN-1998; 98WO-JP07765.

XX 23-JUN-1997; 97JP-0165437.

XX (TAKE) TAKEDA CHEM IND LTD.

XX Fujii R, Hinuma S, Kawamata Y, Matsumoto H;

XX WPI; 1999-105614/09.

XX Use of G protein-coupled receptor ligands - for modulating prolactin
 PT secretion or placental function, e.g. for treating menopausal
 PT syndrome, tumours, autoimmune disease or abnormal pregnancy
 PS
 XX
 PS Claim 3; Page 166; 241pp; English.

CC The present sequence represents a human type ligand fragment. It
 CC is used in the course of the invention. The specification describes
 CC an agent for modulating prolactin secretion which comprises a
 CC ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)
 CC protein. The agents for promoting prolactin secretion can be used for
 CC treating or preventing hypovarianism, gonocyst cacosgenesis, menopausal
 CC syndrome, euthyroid or hypometabolism. They can be used for promoting
 CC lactation in a domestic mammal and as an aphrodisiac. The agents for
 CC inhibiting prolactin secretion can be used for treating or preventing
 CC pituitary adenomatosis, brain tumour, emmenorrhoea, galactorrhoea,
 CC prolactinoma, infertility, impotence, amenorrhoea, galactorrhoea,
 CC acromegaly, Chlari-Frommel syndrome, Argonz-del Castillo syndrome,
 CC Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dyszoospermia.
 CC The inhibitory agents can also be used as contraceptives. The agents for
 CC modulating placental function can be used for treating or preventing
 CC chorioncarcinoma, hydralid mole, interruption mole, abortion, unfertility fetus,
 CC abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.

XX Sequence 20 AA:

Query Match 100.0%; Score 113; DB 20; Length 20;
 Best Local Similarity 100.0%; Pred. No. 3.5e-12;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TPDIINPAWYASRGIRPGRF 20
 DB 1 tpdiinpaawyasrgirpgrf 20

RESULT 3
 B10365
 ID B10365 standard; peptide: 20 AA.

XX B10365;

XX 24-NOV-2000 (first entry)

DE Human oxytocin secretion promoting peptide SEQ ID NO: 35.

XX Human; oxytocin secretion promoter; G protein-coupled receptor protein;
 KW treatment; disease; pain; acute bleeding; uterine recovery failure; cow;
 KW caesarean section; artificial fertilization; galactostasis; goat; pig;
 KW veterinary medicine; milk production.

XX Homo sapiens.

XX W0200038704-A1.

XX 06-JUL-2000.

XX 22-DEC-1999; 99WO-JP07199.

XX 25-DEC-1998; 98JP-0369565.

XX (TAKE) TAKEDA CHEM IND LTD.

XX Matsumoto H, Kitada C, Hinuma S;

XX WPI; 2000-452298/39.

XX Physiologically-active polypeptide recognized as ligand by G
 PT protein-coupled receptor protein, for promoting secretion of oxytocin,
 PT as drugs for diseases relating to oxytocin secretion and in veterinary
 PT medicine

PS Disclosure: Page 63; 72pp; Japanese.
XX
CC This invention describes a novel oxytocin secretion-regulating agent
CC which contains a ligand peptide or its salt for the G protein-coupled
CC receptor protein. It is useful in the form of drugs for ameliorating,
CC preventing and treating diseases relating to oxytocin secretion e.g.,
CC weak pains and atonic bleeding, before and after expulsion of placenta,
CC uterine recovery failure, caesarean section, stoppage of artificial
CC fertilization or galactostasis and is also applicable in veterinary
CC medicine for promoting milk production in cow, goat and pig. This
CC sequence represents a human peptide which acts as an oxytocin secretion
CC promoter.
XX
SQ Sequence 20 AA:

Query Match 100.0%; Score 113; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.5e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWASRGIRPVGRF 20
Db 1 tpdinpaawyarsgrirpvgrf 20
|||||

RESULT 4
Y49294
ID Y49294 standard; peptide: 20 AA.
XX
AC Y49294;
XX
XX 22-FEB-2000 (first entry)
XX
DE 19p2 ligand peptide fragment.
XX
XX Monoclonal antibody; 19p2 ligand; diagnosis; prolactin secretion;
XX pituitary; regulatory mechanism; central nervous system; pancreatic.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX FT Modified-site 20
XX FT /note= "C-terminal amide"
XX
XX WO9960112-A1.
XX
XX 25-NOV-1999.
XX
XX 20-MAY-1999; 99WO-JP02650.
XX
XX 21-MAY-1998; 98JP-0140293.
XX
XX (TAKE) TAKEDA CHEM IND LTD.
XX
XX Matsumoto H, Kitada C, Hinuma S;
XX
XX WPI: 2000-039381/03.
XX
XX New monoclonal antibodies, useful in diagnosis, as drugs and in
XX studying diseases related to ligand abnormality
XX
XX Disclosure: Page 26; 73pp; Japanese.
XX
XX The invention provides a monoclonal antibody which has a specific
XX reaction with the part peptide of the C-terminal of 19p2 ligand or its
XX derivative. The antibodies can be used in diagnosis or to treat or
XX prevent diseases associated with abnormality in the pituitary function
XX regulatory mechanism (e.g. promotion of prolactin secretion), central
XX nervous regulatory mechanism, and pancreatic function regulatory
XX mechanism. The antibody-based immunoassay can also be applied in
XX clarifying the physiological functions of the ligand and its derivative.
XX Sequences Y49290-302 represent peptide fragments of the 19p2 ligand.

SQ Sequence 20 AA:

Query Match 100.0%; Score 113; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.5e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWASRGIRPVGRF 20
Db 1 tpdinpaawyarsgrirpvgrf 20
|||||

RESULT 5
W31395
ID W31395 standard; Peptide; 21 AA.
XX
AC W31395;
XX
XX 06-APR-1998 (first entry)
XX
DE Human type G protein-coupled receptor ligand fragment 5.
XX
XX G protein-coupled receptor; ligand binding; pharmaceutical;
XX modulator; pituitary; central nervous system; pancreas; prophylactic;
XX therapeutic agent.
XX
XX Homo sapiens.
XX
XX W09724436-A2.
XX
XX 10-JUL-1997.
XX
XX 26-DEC-1996; 96WO-JP03821.
XX
XX 18-SEP-1996; 96JP-0246573.
XX
XX 28-DEC-1995; 95JP-0343371.
XX
XX 15-MAR-1996; 96JP-0059419.
XX
XX 12-AUG-1996; 96JP-0211805.
XX
XX (TAKE) TAKEDA CHEM IND LTD.
XX
XX Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;
XX Kawamata Y, Kitada C;
XX
XX WPI: 1997-363672/33.
XX
XX N-PSDB: V02432.
XX
XX Ligand peptide for G protein-coupled receptor - acts by modulating
XX function in the central nervous system, pancreas and pituitary gland
XX
XX Claim 2; Page 186; 258pp; English.
XX
XX This sequence represents a peptide fragment from a novel human type
XX ligand polypeptide corresponding to amino acid residues 34 to 54 of the
XX sequence represented in W31390 and is used in an assay to monitor ligand
XX binding to the G protein-coupled receptor protein. Pharmaceutical
XX compositions containing this ligand may be used as a pituitary function
XX modulator. A central nervous system modulator or a pancreatic function
XX modulator. This ligand could have specific applications as a prophylactic
XX or therapeutic agent for dementia, depression, hyperkinetic syndrome,
XX disturbance of consciousness, anxiety syndrome, schizophrenia, trauma,
XX growth hormone secretory disease, hyper- and polyphagia, hyperlactaemia,
XX hypercholesterolaemia, hyperglycaemia, hyperprolactinaemia, diabetes,
XX rheumatoid arthritis, spinal injury, transient brain ischaemia, asthma,
XX cancer, pancreatitis, renal disease, Turner's syndrome, neurosis, epilepsy,
XX amyotrophic lateral sclerosis, acute myocardial infarction, infertility,
XX spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis,
XX osteoporosis and/or oligosaccharia. Assays can also be developed to screen
XX compounds which are capable of altering the binding activity of the
XX ligand affecting activation of the G protein-coupled receptor protein.
XX
XX Sequence 21 AA:

Query Match 100.0%; Score 113; DB 18; Length 21;
 Best Local Similarity 100.0%; Pred. No. 3,7e-12;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TPDINPAMYASRGIRPVGRF 20
 |||||
 Db 1 tpdinpawyasrgirpvgf 20

RESULT 6

B10366 B10366 standard; peptide; 21 AA.

AC B10366;

DT 24-NOV-2000 (first entry)

DE Human oxytocin secretion promoting peptide SEQ ID NO: 36.

KW Human: oxytocin secretion promoter; G protein-coupled receptor protein;
 KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;
 KW caesarean section; artificial fertilization; galactostasis; goat; pig;
 KW veterinary medicine; milk production.

OS Homo sapiens.

PN WO200036704-A1.

PD 06-JUL-2000.

PE 22-DEC-1999; 99WO-JP07199.

PR 25-DEC-1998; 98JP-0369585.

PA (TAKE) TAKEDA CHEM IND LTD.

PI Matsumoto H, Kitada C, Hinuma S;

DR WPI: 2000-452298/39.

PT Physiologically-active polypeptide recognized as ligand by G
 protein-coupled receptor protein, for promoting secretion of oxytocin,
 PT as drugs for diseases relating to oxytocin secretion and in veterinary
 PT medicine

PS Disclosure; Page 63; 72pp; Japanese.

CC This invention describes a novel oxytocin secretion-regulating agent
 CC which contains a ligand peptide or its salt for the G protein-coupled
 CC receptor protein. It is useful in the form of drugs for ameliorating,
 CC preventing and treating diseases relating to oxytocin secretion e.g.,
 CC weak pains and atonic bleeding, before and after expulsion of placenta,
 CC uterine recovery failure, caesarean section, stoppage of artificial
 CC fertilization or galactostasis and is also applicable in veterinary
 CC medicine for promoting milk production in cow, goat and pig. This
 CC sequence represents a human peptide which acts as an oxytocin secretion
 CC promoter.

CC Sequence 21 AA;

Query Match 100.0%; Score 113; DB 21; Length 21;

Best Local Similarity 100.0%; Pred. No. 3,7e-12;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TPDINPAMYASRGIRPVGRF 20
 |||||
 Db 1 tpdinpawyasrgirpvgf 20

RESULT 7

W31396

ID W31396 standard; peptide; 22 AA.

AC W31396;

DT 06-APR-1998 (first entry)

DE Human type G protein-coupled receptor ligand fragment 6.

KW G protein-coupled receptor; ligand binding; pharmacological;
 KW modulator; pituitary; central nervous system; pancreas; prolactin;
 KW therapeutic agent.

OS Homo sapiens.

PN WO9724436-A2.

PD 10-JUL-1997.

PE 26-DEC-1996; 96WO-JP03821.

PR 18-SEP-1996; 96JP-0246573.

PR 28-DEC-1995; 95JP-0343371.

PR 15-MAR-1996; 96JP-0059419.

PR 12-AUG-1996; 96JP-0211805.

PA (TAKE) TAKEDA CHEM IND LTD.

PI Fuji R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;

DR Kawamata Y, Kitada C;

DR WPI: 1997-363672/33.

DR N-PSDB: V02433.

PT Ligand peptide for G protein-coupled receptor - acts by modulating

PT function in the central nervous system, pancreas and pituitary gland

PS Claim 2; Page 186; 258pp; English.

CC This sequence represents a peptide fragment from a novel human type
 CC ligand polypeptide corresponding to amino acid residues 34 to 55 of the
 CC sequence represented in W31396 and is used in an assay to monitor ligand
 CC binding to the G protein-coupled receptor protein. Pharmaceutical
 CC compositions containing this ligand may be used as a pituitary function
 CC modulator, a central nervous system modulator or a pancreatic function
 CC modulator. This ligand could have specific applications as a prolactin
 CC or therapeutic agent for dementia, depression, hyperkinetic syndrome,
 CC disturbance of consciousness, anxiety syndrome, schizophrenia, trauma,
 CC growth hormone secretory disease, hyper- and polyphagia, hyperlipidaemia,
 CC hypercholesterolaemia, hyperglycaemia, hyperprolactinaemia, diabetes,
 CC cancer, pancreatitis, renal disease, Turner's syndrome, neurosis, asthma,
 CC rheumatoid arthritis, spinal injury, transient brain ischaemia, epilepsy,
 CC amyotrophic lateral sclerosis, acute myocardial infarction, infertility,
 CC spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis,
 CC osteoporosis and/or oligosaccharia. Assays can also be developed to screen
 CC compounds which are capable of altering the binding activity of the
 CC ligand affecting activation of the G protein-coupled receptor protein.

CC Sequence 22 AA;

Query Match 100.0%; Score 113; DB 18; Length 22;

Best Local Similarity 100.0%; Pred. No. 3,9e-12;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TPDINPAMYASRGIRPVGRF 20
 |||||
 Db 1 tpdinpawyasrgirpvgf 20

RESULT 8

B10367 B10367 standard; peptide; 22 AA.

XX

KW Irruption mole; abortion; unthrifty fetus; abnormal saccharometabolism;
 KW abnormal lipidmetabolism; oxytocia.
 XX Homo sapiens.
 OS KO9858962-A1.
 PN 30-DEC-1998.
 PD 22-JUN-1998; 98MO-JP02765.
 PR 23-JUN-1997; 97JP-0165437.
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX Fuji R, Hinuma S, Kawamata Y, Matsumoto H;
 PI WPI, 1999-105614/09.
 XX Use of G protein-coupled receptor ligands - for modulating prolactin
 PT secretion or placental function, e.g. for treating menopausal
 PT syndrome, tumours, autoimmune disease or abnormal pregnancy
 PS Claim 3; Page 159; 24JPP: English.
 CC The present sequence represents a human type ligand fragment. It
 CC is used in the course of the invention. The specification describes
 CC an agent for modulating prolactin secretion which comprises a
 CC 15 and polypeptide or a salt, for a G protein-coupled receptor (GPCR)
 CC protein. The agents for promoting prolactin secretion can be used for
 CC treating or preventing hypometabolism, gonocyst cacosgenesis, menopausal
 CC syndrome, euthyroid or hypometabolism. They can be used for promoting
 CC lactation in a domestic mammal and as an aphrodisiac. The agents for
 CC inhibiting prolactin secretion can be used for treating or preventing
 CC pituitary adenomatosis, brain tumour, emmeniopathy, autoimmune disease,
 CC prolactinoma, infertility, impotence, amenorrhea, galactorrhea,
 CC acromegaly, Chiari-Frommel syndrome, Argonz-del Castillo syndrome,
 CC Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dyszoospermia.
 CC The inhibitory agents can also be used as contraceptives. The agents for
 CC modulating placental function can be used for treating or preventing
 CC chorionicarcoma, hydatid mole, Irruption mole, abortion, unthrifty fetus,
 CC abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.
 XX Sequence 31 AA;
 SQ
 Query Match 100.0%; Score 113; DB 20; Length 31;
 Best Local Similarity 100.0%; Pred. No. 5.8e-12;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TPDINPAWASRGIRPVGRF 20
 DB 12 tpdinpawasrglrpvgrf 31
 RESULT 11
 ID W87615 standard; Peptide; 31 AA.
 AC W87615;
 DE 29-MAR-1999 (first entry)
 XX Human 19P2 ligand.
 KW 19P2 ligand; G protein coupled receptor; pituitary;
 KW prolactin releasing peptide; human; dementia; breast cancer;
 therapy.
 OS Homo sapiens.
 XX EP887417-A2.
 PN

PD 30-DEC-1998.
 XX 25-JUN-1998; 98EP-0111725.
 PR 27-JUN-1997; 97JP-0172118.
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX Moriya T, Nishimura O, Suenaga M, Tanaka Y;
 PI WPI, 1999-047884/05.
 XX Producing a 19P2 pituitary G protein receptor ligand - by cleavage
 PT of a fusion protein, useful for preventing and treating dementia,
 PT breast cancer, renal failure and autoimmune disease
 PS Claim 5; Page 35; 56PP; English.
 CC This is the amino acid sequence of the human pituitary G
 CC protein-coupled receptor ligand 19P2L. A method suitable for
 CC commercial high-level production of 19P2L comprises expressing
 CC the ligand in host cells as a recombinant fusion protein e.g. with
 CC human basic fibroblast growth factor (see V83796-97) that has
 CC been modified to include an N-terminal cysteine residue. The
 CC ligand is released from the fusion by cyanidation followed by
 CC ammonolysis. 19P2L has prolactin secretion-stimulating and (at
 CC high doses) prolactin secretion-inhibiting properties. It can be
 CC used in the treatment and prevention of various diseases including:
 CC senile dementia, cerebrovascular dementia, and dementia associated
 CC with: genealogical disorders (e.g. Alzheimer's disease, Parkinson's
 CC disease, Pick's disease, Huntington's disease), infectious diseases
 CC (e.g. Creutzfeldt-Jakob's), endocrine or metabolic disease or
 CC toxicoosis (e.g. hypothyroidism, vitamin B12 deficiency, alcoholism,
 CC intoxication by drugs, metal and organic compounds), tumorigenic
 CC diseases (e.g. brain tumour), traumatic diseases (e.g. chronic
 CC subarachnoidal haemorrhage, and other types of dementia, depression,
 CC hyperactive child syndrome (microcephalopathy) and disturbance of
 CC consciousness. It is also useful for prevention and treatment of
 CC diseases associated with prolactin hypo and hypersecretion
 CC respectively, including: hyperprolactinaemia, pituitary adenoma,
 CC breast cancer, infertility, impotence and autoimmune disease
 CC (hyperserecton disorders), and seminal vesicle hypoplasia,
 CC osteoporosis, menopausal syndrome and renal failure (hyposecretion
 CC disorders). The 19P2 polypeptide/amide is also useful as a test
 CC reagent for study of the prolactin secretory function or as a
 CC lactagogue in mammalian farm animals.
 XX Sequence 31 AA;
 SQ
 Query Match 100.0%; Score 113; DB 20; Length 31;
 Best Local Similarity 100.0%; Pred. No. 5.8e-12;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TPDINPAWASRGIRPVGRF 20
 DB 12 tpdinpawasrglrpvgrf 31
 RESULT 12
 ID B10362 standard; Peptide; 31 AA.
 AC B10362;
 DE 24-NOV-2000 (first entry)
 XX Human oxytocin secretion promoting peptide SEQ ID NO: 32.
 KW Human; oxytocin secretion promoter; G protein-coupled receptor protein;
 KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;
 KW caesarean section; artificial fertilization; galactostasis; goat; pig;
 KW veterinary medicine; milk production.

```

XX OS Homo sapiens.
XX PM WO200038704-A1.
XX PD 06-JUL-2000.
XX PE 22-DEC-1999; 99WO-JP07199.
XX PR 25-DEC-1998; 98JP-0369585.
XX PA (TAKE ) TAKEDA CHEM IND LTD.
XX PI Matsumoto H, Kitada C, Hinuma S;
XX DR WPI: 2000-452298/39.
XX PT Physiologically-active polypeptide recognized as ligand by G
XX PT protein-coupled receptor protein, for promoting secretion of oxytocin,
XX PT as drugs for diseases relating to oxytocin secretion and in veterinary
XX PT medicine.
XX PS Disclosure; Page 62, 72pp; Japanese.
XX CC This invention describes a novel oxytocin secretion-regulating agent
XX CC which contains a ligand peptide or its salt for the G protein-coupled
XX CC receptor protein. It is useful in the form of drugs for ameliorating,
XX CC preventing and treating diseases relating to oxytocin secretion e.g.,
XX CC weak pains and atonic bleeding, before and after explislon of placenta,
XX CC uterine recovery failure, caesarean section, stoppage of artificial
XX CC fertilization or galactostasis and is also applicable in veterinary
XX CC medicine for promoting milk production in cow, goat and pig. This
XX CC sequence represents a human peptide which acts as an oxytocin secretion
XX CC promoter.
XX SQ Sequence 31 AA:

Query Match 100.0%; Score 113; DB 21; Length 31;
Best Local Similarity 100.0%; Pred. No. 5.8e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYASRGIRPVGRF 20
   ||||||||||||||||
DB 12 tpdinpaawyaaargirpvgrf 31

RESULT 13
Y49291
ID Y49291 standard; peptide; 31 AA.
XX AC Y49291;
XX DT 22-FEB-2000 (first entry)
XX DE 19p2 ligand peptide fragment.
XX KM Monoclonal antibody; 19p2 ligand; diagnosis; prolactin secretion;
XX KM pituitary; regulatory mechanism; central nervous system; pancreatic.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Modified-site 31 /note="C-terminal amide"
XX PN WO9960112-A1.
XX PD 25-NOV-1999.
XX PF 20-MAY-1999; 99WO-JP02650.
XX PR 21-MAY-1998; 98JP-0140293.

```

```

XX PA (TAKE ) TAKEDA CHEM IND LTD.
XX PM Matsumoto H, Kitada C, Hinuma S;
XX PD WPI: 2000-039381/03.
XX PE New monoclonal antibodies, useful in diagnosis, as drugs and in
XX PT studying diseases related to ligand abnormality.
XX PS Disclosure; Page 26; 73pp; Japanese.
XX CC The invention provides a monoclonal antibody which has a specific
XX CC reaction with the part peptide of the C-terminal of 19p2 ligand or its
XX CC derivative. The antibodies can be used in diagnosis or to treat or
XX CC prevent diseases associated with abnormality in the pituitary function
XX CC regulatory mechanism (e.g., promotion of prolactin secretion), central
XX CC nervous regulatory mechanism, and pancreatic function regulatory
XX CC mechanism. The antibody-based immunoassay can also be applied in
XX CC clarifying the physiological functions of the ligand and its derivative.
XX CC Sequences Y49290-302 represent peptide fragments of the 19p2 ligand.
XX SQ Sequence 31 AA:

Query Match 100.0%; Score 113; DB 21; Length 31;
Best Local Similarity 100.0%; Pred. No. 5.8e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPDINPAWYASRGIRPVGRF 20
   ||||||||||||||||
DB 12 tpdinpaawyaaargirpvgrf 31

RESULT 14
W31392
ID W31392 standard; Peptide; 32 AA.
XX AC W31392;
XX DT 06-APR-1998 (first entry)
XX DE Human type G protein-coupled receptor ligand fragment 2.
XX KM G protein-coupled receptor; ligand binding; pharmaceutical;
XX KM modulator; pituitary; central nervous system; pancreas; prophylactic;
XX KW therapeutic agent.
XX OS Homo sapiens.
XX PN WO9724436-A2.
XX PD 10-JUL-1997.
XX PE 26-DEC-1996; 96WO-JP03821.
XX PR 18-SEP-1996; 96JP-0246573.
XX PR 28-DEC-1995; 95JP-0343371.
XX PR 15-MAR-1996; 96JP-0059419.
XX PR 12-AUG-1996; 96JP-0211805.
XX PA (TAKE ) TAKEDA CHEM IND LTD.
XX PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;
XX PI Kawamata Y, Kitada C;
XX DR WPI: 1997-363672/33.
XX DR N-PDB; V02429.
XX PT Ligand peptide for G protein-coupled receptor - acts by modulating
XX PT function in the central nervous system, pancreas and pituitary gland
XX PS Claim 2; Page 185; 258pp; English.

```

XX This sequence represents a peptide fragment from a novel human type
CC ligand polypeptide corresponding to amino acid residues 23 to 54 of the
CC sequence represented in R31390 and is used in an assay to monitor ligand
CC binding to the G protein-coupled receptor protein. Pharmaceutical
CC compositions containing this ligand may be used as a putative function
CC modulator, a central nervous system modulator or a pancreatic function
CC modulator. This ligand could have specific applications as a prophylactic
CC or therapeutic agent for dementia, depression, hyperkinetic syndrome,
CC disturbance of consciousness, anxiety syndrome, schizophrenia, trauma,
CC growth hormone secretory disease, hyper- and polyphagia, hyperlipidaemia,
CC hypercholesterolaemia, hyperglycaemia, hyperprolactinaemia, diabetes,
CC cancer, pancreatitis, renal disease, Turner's syndrome, neurosis, asthma,
CC rheumatoid arthritis, spinal injury, transient brain ischaemia, epilepsy,
CC amyotrophic lateral sclerosis, acute myocardial infarction, infertility,
CC spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis,
CC osteoporosis and/or oligoparesis. Assays can also be developed to screen
CC compounds which are capable of altering the binding activity of the
XX ligand affecting activation of the G protein-coupled receptor protein.
XX

Sequence 32 AA:

SQ **Sequence** **32 AA;**

Query Match	100.0%	Score 113;	DB 18	Length 32;
Best Local Similarity	100.0%	Pred. No. 6,1e-12;		
Matches 20; Conservative	0;	Mismatches	0;	Gaps 0

```

0Y      1  TPDINPAWYASRGIRPVGRF  20
        |||||
Db      12  tpdinpawyasrgirpvgrf  31

```

RESULT
B1.0363

ID B10363 standard; peptide; 32 AA.

AC B10363;

DT 24-NOV-2000 (first entry)

DE Human oxytocin secretion promoting peptide SEQ ID NO: 33.

KM Human; oxytocin secretion promoter; G protein-coupled receptor protein;
KM treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;
KM caesarean section; artificial fertilization; galactostasis; goat; pig;
KM veterinary medicine; milk production.

05 Homo sapiens.

PN WO200038704-A1.

PD 06-JUL-2000

PF 22-DEC-1999; 99WO-JP07199.

PR 25-DEC-1998; 98JP-0369585.

PA (TAKE) TAKEDA CHEM IND LTD.

PI Matsumoto H, Kitada C, Hinuma S;

DR WPI; 2000-452298/39.

PT physiologically-active polypeptide recognized as ligand by G
PT protein-coupled receptor protein, for promoting secretion of oxytocin,
PT as drugs for diseases relating to oxytocin secretion and in veterinary
PT medicine -

PS Disclosure; Page 62; 72pp; Japanese.

CC This invention describes a novel oxytocin secretion-regulating agent
CC which contains a ligand peptide or its salt for the G protein-coupled
CC receptor protein. It is useful in the form of drugs for ameliorating

CC preventing and treating diseases relating to oxytocin secretion e.g.
CC weak pains and atonic bleeding, before and after expulsion of placenta,
CC uterine recovery failure, caesarean section, stoppage of artificial
CC fertilization or galactostasis and is also applicable in veterinary
CC medicine for promoting milk production in cow, goat and pig. This
CC sequence represents a human peptide which acts as an oxytocin secretion
CC promoter.

5Q	Sequence	32 AA
----	----------	-------

Query Match	100.0%;	Score 113;	DB 21;	Length 32;
Best Local Similarity	100.0%;	Pred. No. 6.1e-12;		
Matches 20;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

```

QY      1  TPDINPAWYASRGIRPVGRF  20
        |||
Db      12  LpDINPAWYASRGIRPVGRF  31

```

Search completed: April 17, 2001, 15:38:42
Job time: 349 sec

Tue Apr 17 15:46:19 2001

us-09-446-543a-64.rag

